

Protect us from data protection bill

The British Government's Data Protection Bill is the opposite of what it should be, a libertarian charter. It needs to be amended by the House of Commons.

THE British Government, which last week conceded that the British Medical Association (and the Church of England) has been right these past several months in demanding that the Police and Criminal Evidence Bill should not give the police (on application to a magistrate) the power to breach professional confidences, is going to have to make comparable concessions in its Data Protection Bill, now launched on the House of Commons after an unexpectedly easy passage in the House of Lords. For while the bill has been slightly changed and in some small ways improved, its most obnoxious proposals have become more obviously so. For the bill as it now stands is even less a safeguard of personal liberty than when it seemed (*Nature* 27 January, p.271) a licence for Big Brother.

The defects of this shabby piece of legislation may stem primarily from the cynicism with which it was conceived — the need for something on the statue book that would allow the United Kingdom to sign the European Convention on the regulation of data banks and thus not have British computer bureaux denied access to mainland European markets — but the anomalies that would accumulate if the bill ever became law are too horrendous to contemplate. Ostensibly, the bill is meant to give some measure of protection to people's privacy — that, at least, is what the European Convention asks. But the British bill exempts from the modest protection which it offers all but computerized records, those considered by governments (local as well as central) to serve the purposes of crime prevention, tax evasion and national security. And while individuals will have the right to see a copy of any computer record held under their name, it will be perfectly proper for one who operates such a data bank to decline such a request if there is a danger that third parties, informants say, may in the process be identified. If, for example, *Nature's* correspondence with authors and referees were stored on a computer (which they are not) it would be a sufficient excuse for denying a disappointed author a sight of his referees' opinions that the identity of the referees might in the process become apparent. This flaw is obviously more important, and more directly an infringement of personal liberty, when the information accumulated in data banks is supplied by informants.

This is one rudimentary way in which a Data Protection Bill conforming in spirit as well as by letter with what the European Convention asks should be drafted. Individuals should have an absolute right to see privately operated computer entries under their names. If that entails that informers risk being identified, and that data can then be accumulated only with some difficulty, the result will at least be that more of it will be accurate. But the more serious enormity in the British Government's proposals is that data accumulated in computer banks at the public expense will not be open as of right to the individuals concerned, although the Home Secretary will be able to order whatever access he thinks prudent by other government departments.

DNA now and tomorrow

A report of last week's conference appears at page 651. A second conference on this theme will be held in Boston from 19–21 September. Details of a scholarship scheme to excuse graduate students and some others the payment of a registration fee will be published later.

The consequences could be disastrous, as the example of the "new" disease called Acquired Immune Deficiency Syndrome illustrates all too vividly. Many of the thousand or so people so far affected are male homosexuals, many of whom have a florid history of infection by venereal disease and many of whom have also taken drugs intravenously. The recognition of these common factors predisposing towards the disease is likely to be an important pointer to its causation, perhaps even to some means of treatment. But is it likely that even people suffering from such shabby diseases will give their physicians a truthful history of their case if they suspect that the police will demand access to their records so as to help track down their heroin pusher or even their homosexual partners? British physicians are right to protest at the threat in the bill to the professional confidences (from which lawyers have been exempted). But the threat is not to professional *amour propre* but to the effective prosecution of disease.

The general principle goes further. Few would ask that civil liberty should hamper the effective government of countries such as Britain, which means that government agencies should be allowed (even encouraged) to use computers to do their jobs more efficiently. The Big Brother worry is that government computers programmed initially for different purposes will be used in concert with each other to deal with individuals in ways that are manifestly unjust. So the Data Protection Bill should be amended in the House of Commons to require at least the identification of government data banks — and a reasoned explanation of why some remain closed to those whose names are carried. And government departments' computers should be forbidden to communicate with each other except under the supervision of an independent official with the interests of individuals first in mind.

Looking for a lead

The British Government has made a mistake in its quick embrace of a proposal to ban lead in petrol.

THAT one man's meat is usually another's poison (and vice versa) is not well understood. But what happens when one man (or woman) and the other are the same? That is the dilemma underlying the study carried out in Britain (and published last week, see page 643) by the Royal Commission on Environmental Pollution. At least where adults are concerned, lead pollution of the environment is partly, even principally, a by-product of the beneficial use of this potentially toxic heavy metal; it helps to get them from one place to another economically, without the high-compression engines of their motor-cars disintegrating (or melting).

The fact that they (or more importantly) their children may be damaged as a consequence is a hazard with a different time scale, so long (compared with the inconvenience of a broken-down motor-car) as to be forgotten. The commission is nevertheless right to conclude that there is too much lead in the environment and in people's blood, that switching to lead-free petrol is the only practicable means in sight of reducing lead pollution quickly (which means a decade from now) and that the costs involved should therefore be shouldered stoically. The commission is right, within its terms of reference, to call for this action by the government. The government has however been too hasty in saying that it will bite the toughest bullet — and ban lead from petrol — while



reserving its position on other proposals such as the need to reduce the concentration of lead in the working environment or in a third of Britain's drinking-water. An earlier government's hasty acceptance of the Robbins report on higher education may yet haunt the present Prime Minister, her heirs and her political assignees.

Both parties, the commission and the government, are in cleft sticks. The commission, fully aware that only practicable recommendations will carry weight and properly alarmed that many British residents have enough lead in their blood that overt toxicity is within sight, proposes that lead pollution should be reduced in the course of the next ten years or so by banning lead in petrol. The government, especially aware at this stage in its constitutional term of office that the toleration of pollution must invite electoral unpopularity (to say the least), accepts this drastic measure as a must, but specifically reserves its position on what might be equally valuable strategies such as an attempt to understand what happens in airborne dust (can it contain so much lead, can children consume so much but adults so little?). Neither the oil companies nor the motor-car manufacturers will suffer, for their extra costs will in due course be passed onto their customers, the car-driving lead-breathing tax-paying electors whose interests are meant to be safeguarded by the new policies.

The British Government's decision to embrace the commission's recommendation at the very hour of its publication would be more persuasive if the reasons for the peculiar accumulation of lead in the blood of children were better understood. Naturally, and properly, students of lead pollution have paid particular attention to the apparently rapid accumulation of lead in the bodies of people aged two or three, and to the possibility (even now unconfirmed, however) that their intelligence may as a consequence be impaired. To be truthful, the British commission's report on lead pollution, a model of how public documents should deal with technical subjects, is no more able than its predecessors on this subject to account for the accumulation of lead in young children. Like previous deliberative bodies, the commission supposes that the difference between young children and adults is that the former suck their fingers more assiduously, ingesting more lead-bearing dust in the process. The problem of airborne dust deserves the attention the commission urges, but there is nothing in its report to prove that the relatively higher concentrations of lead in the blood of children are signs that children are more likely to be permanently damaged. It could be that they are more efficient at mobilizing lead from functional organs and of excreting it. None of this would argue against an attempt to reduce lead pollution, even by banning lead additives in petrol, but a measure of deliberation would be seemly.

Even now, two other courses of action than that adopted by the British Government would be wise. For most people in Britain, lead in diet (including piped water) is the chief source of lead contamination, with urban and industrial dust for many people a close second (and for some a first). Dietary lead is not easily controlled, but the lead composition of dust could be tackled within the framework of extant regulations (which, to its credit, the commission recommends). But especially if dust is the chief source of lead contamination in infancy, should not the battle against lead in dust deserve equal importance with that against lead in petrol? And should there not be more public money for research and development on internal combustion engines, not just for the sake of obviating the need for lead additives but in the hope of working an obviously emergent market?

For whatever the hazards of body-borne lead, there is no doubt that lead has now become permanently unpopular. The royal commission's case for reducing the general contamination by lead is strong, and cannot be gainsaid, but it has been shamefully and dishonestly exaggerated by the anti-lead lobby. The maddening but plain truth is that attempts to correlate childhood IQ with blood-lead concentration have so far failed (which is not to say that they should not have been made). Unfortunately, this seems not to have dissuaded the lobbyists from asserting that white — or rather, grey — is black. Now the British Government has uncharacteristically thought it prudent to listen to what the lobbyists have been saying. □

Another team player?

Against the odds, the US Senate has confirmed Kenneth Adelman as arms control director.

THERE can be no disguising the magnitude of President Reagan's personal victory last week in persuading the Senate to brush aside the recommendation of its own Foreign Relations Committee by confirming Kenneth L. Adelman as the new director of the Arms Control and Disarmament Agency (ACDA). The victory was unexpected and, with a vote of 57 to 42, entirely convincing. Those who laboured long to block the nomination, notably Senators Alan Cranston of California and Paul Tsongas of Massachusetts, have been quick to describe the outcome as a setback for arms control that will delight the Soviet Union and confound NATO. They may just be wrong.

The original objections to Mr Adelman's nomination centred, correctly enough, on his personal qualities and experience — a test on which it would be hard for any 36-year-old to measure up to the standards set by Mr Eugene Rostow, whom President Reagan dismissed as director of the agency in January. Mr Adelman's cause was further muddled, during the prolonged con-



firmation hearings, by suggestions that he distrusted the whole notion of arms control and that he deliberately misled his inquisitors during questions about proposed personnel changes at ACDA. But these considerations, however weighty, later became subsumed in the larger question of whether the Reagan Administration itself is committed to a serious and achievable arms control policy.

There is of course no reason why the committee should not have raised this larger question. Its members were probably right to conclude that Mr Adelman lacked the stature and indeed the will to become anything other than the loyal instrument of an already flawed arms control philosophy. But there are at least two reasons for hope that, having finished its negotiations with the Senate, the administration will now be able to pursue a more productive strategy in its negotiations with the Soviet Union.

First, Mr Adelman's confirmation will at last lay to rest the misguided belief that a liberal figure should be placed at the head of ACDA regardless of the political complexion of the rest of the administration. The putative independence of the agency has been at best a sham, at worst a device for ensuring that the President and his chief arms control expert are at each other's throat. Whatever his defects, Mr Adelman has the merit of being in close agreement with President Reagan, Mr Shultz and Mr Weinberger, and therefore in a better position to formulate a policy which has the unanimous support of those who wield real power in the arms control arena.

Second, President Reagan has had to pay a considerable political price to secure last week's vote in the Senate. Eight Democrats and several liberal Republicans who were expected to oppose Mr Adelman's confirmation changed their minds after last-minute pleading from the White House. It is to be supposed that they did so only after extracting a number of firm promises that the President will move swiftly to prove to the doubters that the administration has a genuine interest in progress on arms control. By the same token, the appointment of Mr Adelman, a man trusted by the Senate hawks, may ease the administration's task when it tries to sell an arms control package to the likes of Senator Jesse Helms of North Carolina. □

Lead pollution

British Government goes for lead-free gasoline

THE British Government has accepted a recommendation by the Royal Commission on Environmental Pollution that lead should be eliminated from petrol (gasoline). New motor cars will have to be designed for lead-free petrol from the beginning of 1988. In a report on lead pollution published earlier this week (Cmd 8852, HMSO, £9.20), the Royal Commission argues that no purpose would be served by a further restriction of the lead content of motor spirit below the target of $0.15 \mu\text{g}$ per litre fixed for 1986.

The commission thus accepts the opinion of the government and of the motor industry that the present target concentration is the lowest at which the performance of internal combustion engines can be enhanced by lead additives.

The commission's report and the government's prompt acceptance of its chief recommendation are nevertheless not the outright "victory" for the anti-lead lobby that has been claimed by organizations such as the Campaign for Lead-Free Air, otherwise known as Clear. Rather, the report is a balanced (and readable) account of lead pollution in the United Kingdom which confirms that for a substantial part of the British population, petrol is not the most important source of lead pollution.

The case for lead-free petrol stems rather from the commission's concern that the average blood-lead concentration in the British population is as much as a quarter of that at which symptoms of "frank poisoning may occasionally occur". The commission says that the narrowness of this safety margin is "disturbing", and that there is good reason to reduce the exposure of the general population to lead. Lead-free motor spirit offers the quickest and most certain route to that goal.

The Royal Commission on Environmental Pollution, set up in 1970, independent of the government, and constitutionally is only advisory. The proposals on lead are, however, something of an embarrassment for the government, whose own commissioned report on lead, published in 1979, will appear in retrospect to have been somewhat complacent.

Part of the difficulty is that of estimating the exposure of individuals to lead because of the way in which exceptional environmental circumstances can enormously increase people's intake. Thus it is estimated that the consumption of lead in tapwater can vary from $8.5 \mu\text{g}$ a day to roughly 30 times as much ($255 \mu\text{g}$ a day) where plumbing systems are made of lead and the water is such as to assist its solution. By comparison, in a number of representative and extreme models for lead

intake included in the report, the highest adult intake of lead from petrol amounts to $17 \mu\text{g}$ a day.

For adults, dietary intake accounts for an estimated $100 \mu\text{g}$ a day (with the solder on tin cans accounting for 15 per cent of that). The commission draws attention to the increased intake of lead by cigarette smokers, presumably because of the use of lead-based pesticides on tobacco plantations, as well as by those who drink alcohol.

Children are special cases. In the commission's argument, the chief extra source of lead exposure is the intake by children of lead-bearing dust and soil, which may account for as much as $70 \mu\text{g}$ a day in rural areas and $105 \mu\text{g}$ a day in extreme city environments. One of the few gaps in an otherwise persuasive document is the perhaps inevitable lack of data on the ingestion of lead in dust by adults as well as children.

How much will it all cost?

THE Royal Commission's report on lead pollution includes an interesting calculation of the economic cost of different ways of reducing petrol-borne lead contamination. (For those who need it, there is also a clear account of the difference between diesel and petrol-driven engines, and of the causes of engine knock.)

Following earlier studies, the commission has considered several different policies for reducing petrol-borne lead, including one in which lead would be removed from vehicle exhausts by means of filters. The commission accepts the relative costs calculated by the government's 1979 working party (but supposes that their absolute value will roughly have doubled).

On this basis, the net present value of the total costs over twenty years of keeping 0.4 g of lead per litre of petrol but fitting cars with filters is roughly half the cost of switching to lead-free petrol, while annual costs thereafter would be more than halved.

The report describes research carried out in the past few years at the Atomic Energy Research Establishment at Harwell in the past few years which has demonstrated that filters could be fitted physically within existing exhaust systems, and should remove three-quarters of the lead in exhaust gases. Nevertheless the commission comes down against this option on the grounds that filters would be less efficient when in service than when new, and that only lead-free petrol can give a speedy reduction of pollution from petrol-borne lead.

There is, however, very little to choose in

Blood-lead data are sparse, but sufficient to sustain the commission's case that concentrations are uncomfortably high. Surveys carried out among small population groups in different parts of the United Kingdom confirm that concentrations are highest among young children (2–3 years) and decrease with age. Surveys carried out to satisfy a requirement of the European Community in 1979 and 1981 found average adult blood-lead concentrations of $12.8 \mu\text{g dl}^{-1}$ and $11.0 \mu\text{g dl}^{-1}$ in inner cities and suburbs respectively.

On the question whether lead pollution can be shown to have caused behavioural changes (principally IQ deficits) in young children, the commission is agnostic. It says that on present evidence for the relationship between blood-lead and IQ, there is no way of telling whether the relationship is causal or indirect, with both variables dependent on a third, such as social class (see also *Nature* 14 April, p.561).

One of the curious features of the policy now recommended is that the British Government is bound by a European Community directive that requires that the lead concentration in petrol should not be less than 0.15 g l^{-1} . The government's acceptance of the proposal will require that it should urge stricter environmental standards on its European partners. □

cost between unleaded petrol and the policy now decided for 1986, although after a period of twenty years the annual costs of operating the new policy will be more than twice as great (while decreased fuel efficiency in engines with lower compression ratios will increase British fuel consumption by the equivalent of 1.15 million tonnes of crude oil a year).

Part of the commission's argument is that in these circumstances, the incentive for the development of new kinds of internal combustion engines is strong. It asks that the government should encourage the use of diesel engines in smaller vehicles as well as research and development directed towards solutions of the knocking problem.

The most interesting part of the Royal Commission's report is thus the appendix in which it reprints a document supplied by Atlantic Research Associates, an otherwise unidentified organization, in which the relative merits of different kinds of internal combustion engines are compared. The phenomenon of knock in petrol-driven engines occurs at places in the cylinder to which combustion has not had time to travel, but where high octane fuel can be made, so to speak, "to diesel".

More efficient combustion chambers might function without knocking even if the fuel were merely more sensibly distributed, or if the pattern of combustion in the cylinder were more self-consciously designed. The notation that a rough cut of crude oil might fuel suitably designed engines may be an even great challenge. □

US recombinant DNA

Green light for plant field-trials

Washington

THE Recombinant DNA Advisory Committee (RAC) last week relaxed its restrictions on field testing of genetically-engineered plants in anticipation of a flood of individual requests to approve such tests. Researchers will no longer need to apply for a formal exemption from the committee's rules against deliberate release of recombinant organisms; instead they may proceed with the approval of the local Institutional Biosafety Committee (IBC) and a small "plant working group" made up of RAC members.

The new rule still forbids the deliberate release of other organisms containing recombinant DNA, and sets general guidelines for the sorts of plant field tests that will be permitted. The plants must be cultivated crops and belong to a genus that contains no known noxious weeds; the inserted DNA must be well-characterized and contain no harmful genes; and the test field must be physically isolated from other stands of the same crop.

Because virtually nothing is known about the risks of such tests, the committee also ordered that field tests should include procedures for assessing alterations in and spread of the plant genes, and that results should be supplied to the committee's risk assessment programme. The original proposal, as published in the *Federal Register*, would have permitted field tests solely on the authority of the local IBCs. A majority of the committee voted to retain at least some supervision on the federal level, hence the requirement for review by the RAC working group.

The committee also agreed to drop a proposed restriction on the use of antibiotic resistance genes as selectable markers in the genetically-engineered plants. Winston Brill of the University of Wisconsin had argued that "you need selectable markers that don't kill or debilitate the organism" and that to ban resistance genes would stifle plant work. And Anne Vidaver of the University of Nebraska, who was called in as a special consultant, argued that soil organisms already contain many antibiotic resistance genes, so that fears that the plant genes would introduce a new risk of spreading antibiotic resistance to human, animal or plant pathogens were probably unfounded.

As well as removing the burden from RAC of having to review and vote on each request separately, the new rule may actually also cut down the number of requests. Agrigenetics Corporation, in a comment filed with RAC, noted that at least one earlier proposal agreed by RAC had never been followed by field testing, and suggested that the new rules should discourage applications filed for "publicity or political purposes". Hitherto, the automatic publication of requests in the

Federal Register under the previous system guaranteed a measure of free publicity.

At last week's meeting, the committee also took up the question of what its role should be in light of the demise of the President's Commission on Bioethics and a proposal by Representative Albert Gore (Democrat, Tennessee) to establish a new federal commission to scrutinize the application of recombinant DNA in humans.

The principal question is whether the ethical scrutiny performed by the President's Commission or Gore's proposed commission could be incorporated into RAC. "The ethical concerns are quite different from the charge of this group", said Elena Nightingale of the National Academy of Sciences; but conversely, she said, it may be difficult to discuss "technical issues", especially those affecting human genetic engineering, outside an ethical context.

RAC's charge is at present restricted to biosafety issues. The publicity surrounding the case of Martin Cline, the California doctor who in 1980 performed unsuccessful experiments in which recombinant DNA was inserted into human patients,

was at least in part responsible for arousing interest in new federal controls on recombinant DNA activities.

Gore's proposal, to be introduced later this spring, would establish a non-regulatory but independent advisory commission to monitor human genetic engineering. The commission would be made up predominantly of non-scientists, and the focus would be on ethical issues. The proposal is similar to an earlier recommendation of the President's Commission which, in a report last fall, argued that "the time has now come to broaden the area under scrutiny to include issues raised by the intended use of the technique rather than solely the unintended exposure from laboratory experiments".

Henry Miller, the Food and Drug Administration's liaison with RAC, suggested that these proposals may have "missed the mark". He said that FDA will probably regulate the products and processes of human gene therapy, as it does other human treatments, and that other existing mechanisms, including institutional human experimentation committees and RAC itself, ensure adequate scrutiny. "We would submit that arguably it is neither necessary nor sufficient to establish a new regulatory entity", he said.

Stephen Budiansky

Biotechnology in West Germany

Brain drain threatens progress

As in other European countries, there are widespread fears in West Germany that the country might fall behind in the fast moving field of biotechnology. The agreement between Hoechst and Harvard University to cooperate in work done in the United States has therefore caused concern throughout Germany, although decisions by the other two large chemical groups to support research centres in Germany have done something to restore confidence. Bayer has agreed to provide DM3 million over three years to the Max Planck Institute (MPI) at Cologne for research in breeding, while BASF will fund research at Heidelberg for five years at DM1 million per year.

In fact there has been a federal programme on biotechnology since 1979 and last August the then minister at the federal ministry of technology (BMFT), Andreas Von Bülow, proposed a 14 per cent increase in support for this year and a radical review for the future. This task has now fallen to his successor, Heinz Riesenhuber, who promised last month to give high priority to biotechnology research. Nevertheless, the state of biotechnology is still causing concern, as has been shown at the latest meeting of the senate of Max Planck Gesellschaft (MPG) in Stuttgart. Perhaps in response to the federal government's call for closer contacts between research and industry, the MPG president, Professor Reimer Lüst,

has been pressing for more cooperative projects at the Max Planck institutes.

On the other hand, there are already numerous institutes more or less directly involved with genetic engineering and biotechnology, some with industrial support; for example biochemistry in Martinsried, biology and virus research in Tübingen, genetics in Berlin, immunology in Freiburg, experimental medicine and biophysical chemistry in Göttingen, cell biology in Ladenburg and selective breeding at Cologne.

Among the *Länder*, Baden-Württemberg is providing DM30 million for new facilities at the University of Heidelberg; in Berlin there is cooperation between the authorities and the pharmaceutical firm Schering with MPI for molecular genetics. Support from the *Länder* is also being given in Munich and Cologne.

The main concern at the MPI senate meeting was the shortage of suitably qualified researchers. Many have joined the brain drain across the Atlantic and more resources are needed to stop and eventually to reverse the flow. Although these are expected to be provided by the federal ministry in part, it was felt that such funds should not be channelled into large state research centres. Thus biotechnology will do little to help BMFT to solve the problem of what to do with those centres that have outlived their original purpose.

J.S. Dunnett

US centre for advanced materials

Synchrotron source unwelcome

Washington

THE Reagan Administration's proposal to create a National Center for Advanced Materials (NCAM) at Lawrence Berkeley Laboratory (LBL) has received a chilly reception. Nearly 100 materials scientists have written to the House of Representatives Science and Technology Committee criticizing the proposal and the administration's failure to seek the advice of outside experts before unveiling the plan in the President's 1984 budget.

In response to these criticisms, a House subcommittee chaired by Representative Don Fuqua (Democrat, Florida) has cut \$5 million, or about 20 per cent, from the first-year construction budget in order to delay the project and allow for revisions. The subcommittee also found that the administration had "by-passed the preferred review process. . . in its eagerness to move ahead" and ordered that no funds should be set aside for construction until an outside technical review is completed.

The Department of Energy (DoE), which funds LBL and the other national laboratories, has already put together a panel to carry out the review. The chairman is Albert Narath, vice-president for research of Sandia National Laboratory. The other members are drawn from industry, universities and the national laboratories. LBL is also making an effort to smoothe ruffled feathers by sending out 2,000 invitations to a "users' workshop" on the synchrotron Advanced Light Source, the centrepiece of the materials centre.

These belated overtures to the materials research community reflect an apparent miscalculation by George Keyworth, the White House science adviser, who is generally credited with getting the materials centre proposal into the 1984 budget. According to William Brinkman of Bell Laboratories, "this proposal never went through DoE in a conventional way. It went straight to Keyworth's office and was kept under wraps by Keyworth until it went into the budget — without any input from the community."

According to a reliable Washington source, it also went into the budget over the objections of Keyworth's own staff at the Office of Science and Technology Policy.

One of the most outspoken critics of the plan is Rustum Roy, director of the materials research laboratory at Pennsylvania State University. Roy is particularly critical of tying a materials research centre to the new synchrotron source. "It is a perfectly good physics tool, but to the materials community this would be 99th on a list of 100 priorities", he said. "Its impact on the materials field is zero."

Roy said the emphasis should be on synthesis and processing of new materials — as is the Japanese programme. Although the

LBL centre will have one division devoted to synthesis, Roy questioned LBL's expertise. He said a more logical choice would be Oak Ridge National Laboratory.

According to a Fuqua subcommittee staff member, many of the letters from materials scientists also questioned the need for a new synchrotron source. A number also questioned whether materials science is really "big science", centred around big machines, or whether the \$174 million for the LBL centre could be more fruitfully spent on the smaller materials laboratories supported by the National Science Foundation.

The first inkling that outsiders had of the NCAM proposal came last September, when the Solid State Sciences Committee of the National Academy of Sciences (NAS) received details on the storage ring of LBL's planned synchrotron. As concern began to grow about the NCAM plan, LBL took some small steps towards involving outsiders. According to Rob Johnson of LBL, the laboratory invited several members of the academy committee and the Council on Chemical Research to sit in on a DoE review of the proposal, held at the laboratory on 10 January. But the visitors did not offer any formal recommendations,

and, Johnson admits, "at one time we called it a 'review' and people got very upset".

The first complete public presentation of the proposal came on 8 February, a week after the budget was released, when David Shirley, LBL's director, appeared before the NAS committee in Washington.

Johnson said it was difficult to achieve a consensus among materials scientists in part because "it's hard to define a 'materials scientist' ". And he pointed out that unlike high energy physics and nuclear physics, materials science has no standing advisory panel in DoE that could naturally be called upon to review proposals.

In support of LBL's plan, Johnson pointed to a recent National Academy study which concluded that "the growth in the use of synchrotron radiation has been so great" that existing facilities will meet demand only until 1985. And he said that LBL has received "reams of letters from industrial laboratories saying how enthusiastic they are about the NCAM project".

But such arguments fail to impress critics. Roy said it is standard procedure to canvass support from potential users. "Of course if a vice-president for research gets a letter saying, 'we are going to build a light source, will you use it?' — what is he going to say, 'no'? What they're saying is that if the government puts \$140 million into it and we will be charged \$100 an hour [to use it] we will like it." **Stephen Budiansky**

Power from oil shale

Desert plant

Jerusalem

OIL shale, Israel's only major fossil fuel source, will be exploited by a new company set up with 70 per cent government participation. A wide spectrum conversion plant has been developed at the Casali Institute of Applied Chemistry of the Hebrew University of Jerusalem, and small-scale commercial operation is due to begin this month.

This seems, at first, surprising, since although liquid fuel can be extracted from shales it costs some \$60 per barrel, more than twice the current spot market price. The Israeli Ministry of Energy and Infrastructure is, however, willing to support the project, not only as a insurance against the time when world oil stocks begin to run dry, but also because the fluidized bed reactor provides a highly flexible gasification-liquefaction process which can be used either as a source of liquid fuel or as a power source for steam generation.

This flexibility, says Dr Ze'ev Aizenshtat, the team leader, could be an important selling point. The reactor can be used with any particle size from dust to large lumps and any grade of fuel from lowest grade shale to best coal. Although designed for Israeli shales, therefore, it could equally well be used to process the large quantities of low-grade shales in the United States if the Americans decided to

exploit these resources.

The development of solid fuel technology became a priority in Israel after the 1974 oil crisis, and during the past five years, with the return of the Alma oil fields to Egypt under the Camp David Agreement and the loss of regular oil supplies from Iran after the Ayatollah's revolution, it became even more imperative to find an alternative to oil for generating electricity. The use of imported coal, however, as in the new Hadera power station, is seen only as a stopgap, and one major advantage of the Casali Institute's reactor is that it could be incorporated into existing solid-fuel power stations with only relatively minor modifications.

Professor Elata, chief scientist at the Ministry of Energy and Infrastructure, says that the shale reactor is only one of a number of energy options that include several specifically Israeli concepts such as solar-pond powered turbines and hydroelectricity generated by the proposed Mediterranean-Dead Sea watercourse. One important spin-off of the route decided for the latter (through the Negev, south of Be'er Sheva) is that it will bring large quantities of seawater close to the site of the shales. Although the Casali reactor has been designed to minimize water consumption, a considerable amount of water would still be necessary for any large-scale use. The construction of the waterway could therefore be linked with the building of power stations to process the shales *in situ*. **Vera Rich**

Remote sensing satellites

Landsat at centre of storm

Washington

THE Reagan Administration's plan to include the government's weather satellites in a possible sale of the Landsat remote sensing satellites to private business has run into trouble on Capitol Hill. Two panels of the House of Representatives Committee on Science and Technology last week launched hearings on the proposal and expressed "serious concern" about how the decision had been taken.

Representative James Scheuer, the New York Democrat who is chairman of the hearings, said the decision to include the weather satellites to make the purchase of the Landsat system more alluring appeared to have been taken despite a series of government studies concluding that the weather satellites should not be turned over to the private sector. In April 1982 the same conclusion was endorsed by the cabinet council on commerce and trade.

"Eight months later, that decision was apparently reversed, even though neither the cabinet council nor any of its work groups had met in the interval. In fact, if one examines the advice which the administration received in the interval, it is almost inconceivable to me that the decision could have been reversed", he commented.

Committee investigators trying to reconstruct the steps leading up to the decision to include the weather satellites in a sale have discovered that one potential buyer, Comsat Corporation, was allowed to take part in executive branch discussions of the issue, to which neither Congress nor other interested companies were invited. And Mr Malcolm Baldrige, the Secretary of Commerce, revealed during last week's hearings that his deputy secretary, Mr Guy Fiske, and several assistants had been taken off the Landsat negotiations after Comsat offered a job to Fiske while talks with the department were still in progress.

Disclosure of the close ties between the commerce department and Comsat makes it unlikely that Mr Baldrige will be able to press ahead rapidly with a search for potential Landsat and weather satellite buyers. Testifying for the first time since President Reagan announced the controversial proposal last month, the commerce secretary insisted that only by putting the Landsat system in the hands of a competent company with marketing expertise could the United States beat off competition from foreign countries seeking customers for data produced by their own satellites.

France would be launching the first in a series of four land sensing satellites next year, Mr Baldrige said. Japan, Canada, Germany, the European Space Agency, India and the Soviet Union are also planning remote sensing programmes. "It would be ironic if these countries, building on the skills developed in the United States,

established the kind of market that the National Oceanic and Atmospheric Administration (NOAA) can't and turned us into a net importer of satellite data products and services", he said.

But even Mr Baldrige had to admit that an acceptable buyer for Landsat might not turn up, because of all the military and diplomatic strings which the government would have to attach to a sale. Diplomatically, the United States would have to ensure that the proprietary claims of an owner would not undermine the "open skies" policy which made Landsat data available to the many foreign countries which depend on it for vital agricultural information.

The military strings would have to be knotted even tighter. The Department of Defense, in interagency talks on the sale, is insisting that some way be found to return the satellites to government control in an emergency. More difficult still, it wants to be able to screen new sensors attached to the satellites to ensure that they cannot detect information the Pentagon wants to keep secret.

According to Mr Baldrige, the only way to discover whether private companies could meet these conditions is to write them into a specification for bids and test the market. Members of the science committee, having learned that Comsat officials have met defence department experts to examine the military problems, have questioned whether a request for bids will give all companies an equal chance.

The administration argues that many organizations use Landsat data — mineral and energy companies, hydrologists, timber companies, fibre and food crop interests — but that NOAA is not able to exploit this potentially lucrative market as efficiently as a private company.

Many scientists, including officials within NOAA, believe that selling Landsat now would be premature since most of its custom comes from the federal government and it is likely to be six to ten years before the satellites could break even as a commercial venture. The administration concedes that buying Landsat might not appeal to the private sector unless the deal was sweetened by the inclusion of the weather satellites and a guarantee that the government would purchase the data the systems produced.

Representative Scheuer has made no secret of his view. He told Mr Baldrige last week that he agreed on the importance of an initiative to maintain the United States' dwindling technological lead in the satellite area. But he added: "I do not agree, nor do any of the experts who have looked at this problem, that we should sell a weather satellite system that is working well in order to cross-subsidize our land remote sensing capability".

Peter David

US space policy

Still behind closed doors?

Washington

DISCONTENT with US machinery for making policy on space has now provoked some in Congress to try to steal the initiative from the administration. Senator Ernest Hollings of South Carolina, the latest presidential hopeful to join the race for the Democratic nomination, has introduced a bill calling for the creation of a national space commission to overhaul space policy-making and wrest a degree of control from the National Security Council.

Introducing the bill last month, Senator Hollings said that the important decisions on space now took place behind closed doors and were dominated by the president's Space Interagency Group (SIG), an arm of the National Security Council.

The bill, the National Commission on Space Act, would set up a panel of 15 presidential nominees drawn from government, science and the private sector with a one-year mandate to review military and civilian space programmes and recommend legislative changes. It would be empowered to scrutinize government agencies, call hearings and conduct its own studies.

Supporters of the Hollings bill are taking as their text a report on civilian space policy prepared last year by the Office of Technology Assessment (OTA) in response to a request by the Senate Committee on Commerce, Science and Transportation, on which Senator Hollings is the ranking Democrat. The report contained harsh criticisms of existing policies, saying that institutions and principles underlying the American space effort have scarcely changed since the National Aeronautics and Space Act of 1958.

According to OTA, some kind of multi-representative forum is urgently needed to answer the big questions facing space policy: whether the United States is committed to exploring space; how it can become a reliable partner in international ventures; the role of government and private industry and whether the National Aeronautics and Space Administration (NASA) should continue to operate the shuttle once it has completed research and development.

OTA has also strongly criticized the composition of the Space Interagency Group under the Reagan Administration. Dr John Gibbons, OTA's director, has told Congress that by cutting out the private sector and excluding agencies such as the Departments of Agriculture and the Interior and the National Science Foundation, SIG had given "unambiguous control" over policy decisions to the military/intelligence community.

OTA's offensive against the present management of space policy is, however,

likely to continue. Another report, highly critical of the performance of the US delegation at last year's Unispace 82 meeting in Vienna, is about to be published and will argue that the failure to develop long-term domestic policy goals for space has damaged Washington's ability to use space technology as a tool of foreign policy. A prepublication draft of this report says the United States approached the UN-sponsored meeting warily and without enough preparation. The delegation's most obvious mistake was to try, on procedural grounds, to block discussion of the militarization of space — a tactic that merely antagonized other participants.

There were other blunders, too. The United States failed to soothe the fears of developing countries that the geostationary orbit was becoming congested by the satellites of the industrialized nations; nor did it allay anxieties that the confusion sur-

rounding the future of Landsat would jeopardize the availability of remote sensing data. OTA warns that these errors will rebound against the United States, which depends on developing countries for the coordination of critical space systems, such as the allocation of frequencies by the International Telecommunications Union.

A call for the creation of a space commission to take a new look at space priorities may have attractions for scientists as well as for politicians. Professor Arden Albee, chief scientist at the Jet Propulsion Laboratory at Pasadena, California, said it was clear that NASA's main focus was on big systems like the shuttle. He added: "Although its charter includes science, science has not received the level of focus it needs. If the commission resulted in a new look at the priorities of the civilian programme it would be helpful."

Peter David

Semiconductors

British venture nears turning-point

INMOS, the government-backed British microchip company, appears confident that it will reach profitability during 1984 if market conditions improve, despite the revelation in last week's annual report that the company made a loss last year of £20.4 million. Sales revenues increased over the year from £2.14 million in 1981 to £13.7 million and are expected to be double this amount next year.

The company now has three very large-scale integrated circuit memory devices in production at its new factory at Newport in Wales. Dr Richard Petritz, Inmos's managing director, said in his report that during 1983 Inmos will market a 64K erasable reprogrammable memory (EPROM) that will be clearly ahead of its competitors in specification and capacity. The company is also working on a new type of microprocessor called a "transputer". The device is said to be based on a "radically different" design approach and the company expects it to meet the demand for massive computing power in fifth generation "intelligent" computers. The transputer will use a new programming language called Occam, now being

evaluated. A company spokesman said that "excitement" about the transputer had been shown in several quarters, including some Japanese institutions.

The company's overall strategy is to develop products of very advanced specification and to hope that market demand follows. Inmos already claims 60 per cent of the world market for its very fast 16K static random-access memories (RAMs) and hopes for similar success with its new 64K dynamic RAMs. Static RAMs, though now a relatively small and specialized market area, are expected to find increasing applications in high-speed signal processing for military equipment and video systems. But Inmos will need to sell large quantities if it is to justify the high capital costs of its advanced fabrication equipment. Independent market opinions are less confident that Inmos will make headway against Japanese and United States competition with its dynamic access RAMs, where price may be more important than high performance.

The British Government's stake in Inmos is through the British Technology Group (BTG), which holds 75 per cent of the issued share capital. BTG's initial £50 million investment was made by the National Enterprise Board in 1978, when the company was set up by the Labour government. This was increased by an extra £15 million in January this year, but BTG, whose own future is being reviewed, is unlikely to be a further source of capital. Inmos says that it had always been foreseen that private sector investment would be sought during 1983-84, and Mr Malcolm Wilcox, a banker, was appointed chairman in January to carry out the task. With the major start-up costs of the Newport factory out of the way, and production on target, Mr Wilcox's job may just be possible.

Tim Beardsley



Cambridge science park

Sowing a seed

THE relationship between the Cambridge Science Park and the University of Cambridge should be strengthened as the result of an interesting grant scheme proposed recently by Trinity College. Trinity, the park's founder and landlord, has announced that it will provide a total of £60,000 over three years to meet up to half the cost of employing two to three research staff for companies on the science park. The researchers would be full employees of the companies but would spend half their time in the university laboratories. This would give the companies access not only to the university equipment and libraries, which are already made available through Trinity, but also to the expertise of academics working on similar projects.

Trinity's involvement is to be purely financial. The researchers would have no academic relationship with Trinity, although it is hoped that they might teach in their university department. The projects would be arranged between company and university, with Trinity backing the most promising ideas.

The scheme has been met with restrained enthusiasm from companies on the science park. David Storey, managing director of LKB Biochrom, one of the earliest arrivals, sees the scheme as a step towards a better relationship with the university. The system is already working well, he says, but



Trinity — flying a flag for research

any further improvements would be good for industry. Storey says LKB could be keener if it were more involved in research.

In the short term, Trinity College stands to gain nothing from the arrangement but the goodwill of the science park and the university, but with only 56 acres of the 124 acre site developed, the investment of 0.5 per cent of the college's annual external revenue in public relations is a sound financial move. Many of the companies on Cambridge Science Park, such as Bethesda and LKB, are British subsidiaries of international concerns whose research and development is based elsewhere. Trinity hopes to tempt more of it to Cambridge.

Melanie Kee

Science and Jewish Orthodoxy

A row in the making

Rehovot

THE Hebrew University in Jerusalem was last month host to a gathering that has brought the subject of Creationism to the fore and stirred up strong emotions. The "First Congress on Inquiries into the Origin of Life and Evolution" was organized by Orthodox Jewish scientists from Israeli institutions of higher education and took place under the auspices of the Israeli Ministry of Education with the support of the Israel National Council for Research and Development and the Israel Academy of Sciences and Humanities.

A charismatic, ultra-Orthodox Jerusalem rabbi called Mordechai Arnon had previously raised the issue of Darwinism in Israeli secular schools, but since he was a comedian before "seeing the light", his initiative was not taken seriously; now, with leading scientists and official bodies involved, the issue cannot be ignored.

Some participants at the congress, while finding fault with the Darwinian and neo-Darwinian approaches, also said that the other doctrines, including that put forward in *Genesis*, are open to question. However, Professor Alvin Radkovsky, of Ben-Gurion University of the Negev in Beer-sheba, felt otherwise. In a talk on miracles, Radkovsky said scientific research had already demonstrated "the extreme

improbability that life originated and developed by evolutionary means. This being the case, there is no other possibility than the divine origin of life." His view was echoed by Dr Duane Gish, a leading Christian creationist from the United States.

Professor Cyril Domb, a physicist who some years ago left Kings College, London to join Israel's Bar-Ilan University, took a different line. He asserted that many aspects of the theory of evolution had been refuted, but not in time to prevent the theory from doing "enormous moral damage to Western society" by encouraging a belief in man's animal nature and in a meaningless Universe. At the same time, he saw no point in portraying the *Genesis* account of events as an alternative scientific theory as its acceptance was based on revealed truth, not on experimental evidence; it was thus above science.

The main concern at the congress was the threat of Darwinism in secular schools, but computer expert L.M. Spetner warned that it had also found its way into religious schools. "If Darwinism is to be in the curriculum at all", he declared, "then it should be only as a means of preparing students to confront it."

Most non-observant Israel scientists disagreed with the views heard at the meeting, but so did many Orthodox scientists. For example, Weizmann Institute

Professor Michel Revel, a leading expert on interferon and as an observant Jew, dismisses the organizers of the conference as a vociferous and unrepresentative group. Revel sees the account of creation in *Genesis*, with its progression from lower to higher forms of life, as being in broad agreement with Darwinism. He is disturbed by the growing intolerance he finds in religious circles as well as by the attempts made to pit religion against science. "When a child is told to accept something on faith rather than on the basis of scientific evidence, he is liable to lose that faith when he finds that science was right after all", Revel warns.

Nechemia Meyers

Soviet-Australian relations

Signs of thaw

Canberra

SOVIET-Australian cultural relations are coming in out of the cold after more than three years. The new Labor Government is likely to abandon the sanctions applied against the Soviet Union after Soviet troops moved into Afghanistan. Since 1979, bilateral scientific cooperation has been suspended, all high-ranking Soviet officials excluded and academic exchanges banned. Relations between the two countries reached a nadir last August when the entire team of Soviet registrants boycotted the international congress of biochemistry held in Perth under the aegis of the International Union of Biochemistry, after two Soviet scientists were refused visas (see *Nature* 299, 97; 1982).

Foreshadowing the change in policy, the Prime Minister, Mr Bob Hawke, said last month that there is a "lack of consistency" in Australian relations with the Soviet Union, with sanctions in some fields but business as usual in the export of wool and meat to the Soviet Union.

One of the casualties of the sanctions has been the bilateral science and technology agreement signed in January 1975. The agreement has as one of its aims "the exchange of scientists, technical specialists and delegations with the aim of carrying out scientific research" and as another "the pursuit of joint research activities". Only last week, an ex-chief of the division of entomology at the Commonwealth Scientific and Industrial Research Organization, Dr D.F. Waterhouse, was made an honorary fellow of the Soviet Academy of Science in recognition of his contributions towards scientific cooperation with the Soviet Union in entomology.

Fortunately, cooperation, scientific and otherwise, has never been suspended in Antarctica. Recently an Australian supply ship making its last run for the year to Hobart later than usual, found itself caught in pack-ice in Prydz Bay. A Soviet icebreaker from a nearby base responded to a call for help, but as it turned out its services were not needed.

Vimala Sarma

Never on Saturday

Rehovot

ORTHODOX Jewish scientists may disagree on evolution, but all agree that the tools of science should be used to facilitate ritual observance of the Sabbath. The problem is particularly acute in Israel, where vital services must be maintained by Jews throughout the year even though Orthodox Jews will not ordinarily so much as turn on a light on the Sabbath.

When is a 'phone not a 'phone?



Thus religious Jews often install timers in their homes that turn their lights off and on at predetermined times on Saturday. Professor William Z. Low, president of the Jerusalem College of Technology has recently pointed out that this approach of continued action combined with "indirect causation" can be used to solve much more difficult problems where essential work must be carried out by the army, police or hospitals. For instance a set of pre-programmed instruction cards could be inserted on the Sabbath into a processor in a nonoperational mode. The processor would be programmed (before Saturday) to self-activate automatically every few seconds, to read the instructions and to respond, thus effecting complicated changes in factory operations.

Special technological answers to the problems of Sabbath observance are also useful to Orthodox Jews in the political sphere. Israel's Ambassador to the United Nations, Professor Yehuda Blum, has been greatly aided by his special Sabbath telephone, which works on the basis of "indirect causation". When it rings and Blum picks it up, he is not automatically connected with the person at the other end of the line; that happens only when an automatic electronic device "sees" that he has lifted the receiver and activates his line.

Nechemia Meyers

Animals' rights and wrongs

SIR — The case for animal rights is really much simpler than your anonymous article (*Nature* 24 March, p.287) would suggest. Whilst animals differ from us in many ways, such as intelligence, strength, language and appearance, these differences are entirely arbitrary and can occur *within* our own species. Animals are like us in one crucial respect — they too can experience pain, fear and stress, and on that basis the National Anti-Vivisection Society believes they should receive the same respect and consideration we would expect for each other. The suffering which is so often inseparable from animal experimentation, and the intensive forms of animal husbandry, hardly approach this ideal.

It is only the enormous implications of such an ethical revolution that complicate the issue. In the same way, logical arguments for the abolition of slavery were clouded by implications of economic collapse.

In the case of animal experiments, numerous powerful influences work to keep the *status quo*. The introduction of potentially highly profitable products is legally dependent on animal tests whilst scientists will be averse to losing one of their traditional "research tools". Nor should we forget the laboratory animal breeders and associated industries.

With such powerful vested interests there has been understandable, though lamentable, lack of government sympathy. But recent moves to link votes with animal protection promise to make this an important political issue and it may well be that governments will be finally persuaded to take steps towards reform. To begin with this could mean eliminating the LD50 test, restricting funds for the use of animals in psychological and behavioural research, and real government support for the development of alternative methodologies. In the case of toxicity testing, at least, the need for humane and more reliable tests has long been apparent.

Yet how much better it would be for government and scientists to take the initiative *now* and respond positively to growing public concern, instead of acting defensively as your article rather negatively suggests. The question of animal rights can no longer be ignored.

ROBERT SHARPE

National Anti-Vivisection Society,
London W1, UK

SIR — Your News and Views comment, "Do laboratory animals have rights?" (*Nature* 24 March, p.287) fails to point out that Professor Peter Singer's notions of animal rights, though popular, have dubious foundations.

(1) Peter Singer ignores the basic distinctiveness of the human experience of pain that underlies morality. While nonhumans experience pain in isolation

from any beliefs, it is a species-specific characteristic of human suffering that it is modified by beliefs concerning its aetiology. A simple example of this is the exclusively human phenomenon of placebos. A more complex example is the different affective reactions people have to similar injuries when they believe they are accidents compared with when they believe they are the products of intention. People also suffer differently a similar physical injury according to whether they have consented for example to medical treatment or have it done against their will. There is no comparable sentience in nonhumans. The involvement of beliefs is central to human experience and our sentient nature. It not only underlies our capacity to conceptualize moral frameworks but the human need for them.

(2) His attack on "speciesism" using analogies that animal experience has with that of retarded and infant humans ignores the importance of generalizing human rights to all classes of human experience. If certain groups of humans were exempted there could be the risk that by accident or design — witness present-day psychiatric practice in the Soviet Union — fully sentient humans would be denied their rights by being misclassified.

(3) For most of this century anti-vivisectionists have concluded, since we cannot predict the beneficial consequences of pure research using vivisection, that no moral obligation to human welfare is effected by restricting it in the interests of animal welfare. However, though on any particular occasion of its being asserted this is no doubt true, retrospectively all such claims until now have been shown to be false by the sequential development of medical science. I cannot invalidate Peter Singer's claims that in restricting vivisection he is not restricting future human welfare, but it is clear that all previous assumptions to this effect have been shown to be erroneous and present-day claims might be equally so. It is reasonable to believe that restricting research in the name of animal rights is to act contrary to our moral obligations to future generations not knowingly to restrict the opportunities they might have to alleviate or prevent human suffering. To have acted on animal rights in the past would have denied many patients today the care that medical progress has produced. Can it be right to deny the future improvement in treatment efficacy consequent to implementing the unproven concept of animal rights?

JOHN R. SKOYLES

London NW3, UK

SIR — The cacophonous neologism "speciesism" is used (*Nature* 24 March, p.287) in a discussion of animal experimentation, quoting Professor Peter Singer. Singer's dicta extend beyond this; he has

been quoted as saying¹: "What are we to do about genuine conflicts of interest like rats biting slum children? . . . the essential point is just that we do see this as a conflict of interests, that we recognize that rats have interests too."

A candidate for the Sierra Club Board of Directors has pushed it a little further in his election statement by saying "All life forms from blue whales to bacteria have an equal right to exist". This raises the question of the abrogation of the rights of intestinal coliforms by the end result of peristalsis. Also, it is doubtful whether pathogenic microorganisms perceive the rights of their hosts. It brings us back to the darwinian notion of the struggle for existence, in which our own species has been perhaps over-successful, but not sufficiently successful as to stop competing with other biota. Obviously the inflicting of needless pain is to be avoided.

THOMAS H. JUKES

Department of Biophysics
and Medical Physics,
University of California, Berkeley, USA

1. *Parity Foundation Newsmagazine* Vol. 1,7 (1981).

Toxin gene cloning

SIR — In considering the applicability of cloned toxin genes to biological warfare, James Larrick's letter (and indeed, the title your editors attached to it, "Beware cloned toxins") suggests that cloned toxins are more dangerous than these toxins produced in their natural hosts (*Nature* 24 February, p.651). This is highly improbable. In fact, it is likely that these cloned toxins will be less dangerous (in *Escherichia coli* K-12, at least) because of the differences in cellular compartmentalization and post-translational processing which must accompany the secretion of these toxic proteins from their natural hosts.

The example cited by Mr Larrick, diphtheria toxin, illustrates this point well. Diphtheria toxin produced by certain (widely available) strains of *Corynebacterium diphtheriae* may produce and secrete up to 500 μ g per millilitre of diphtheria toxin (this is equivalent to about 70,000 human LD₅₀s per litre). One hardly needs a recombinant *E. coli* strain to manufacture this toxin. In addition, the highly toxinogenic strains of *C. diphtheriae* mentioned above are essentially avirulent due to their inability to colonize the throat, pointing to the multifactorial requirements of virulence.

There is a great need for public awareness about the potential dangers and abuses of chemical and biological warfare. The particular letter you chose to print, however, did little to advance this objective.

KEN COLEMAN

Harvard Medical School,
Boston, Massachusetts, USA

Mauna Kea — the best choice

SIR — In his letter on the site for the UK millimetre wave telescope (*Nature* 24 March, p.286) Dr H.A. Gebbie stated that the "project was started in the late 1960s" and he recalled a conversation with Mr J.F. Hosie of the Science Research Council (SRC) to the effect that "funding had been agreed and the delay in building was giving him concern". Until September 1970 I was chairman of the Astronomy Space and Radio Board of SRC and perhaps I may be allowed to place on record the exact sequence of events concerning the millimetre wave project in the late 1960s.

In January 1968 Professor Graham Smith presented a paper to the Astronomy Committee of the board on the potential importance to astronomy of the millimetre region of the waveband. He advised the committee that "advances in millimetre wave astronomy were not obtainable without considerable cost and some excursion into relatively unexplored fields". The committee set up a working group with Professor Smith as chairman to review the "rewards and difficulties" of a millimetre wave programme. This group began its work in February 1969 and at a meeting of the SRC committee in February 1970, Professor Smith presented his report. The main recommendation was that efforts should be concentrated on providing the best possible instrument for observation in the 8 to 10 mm range and that to this end "a study of a 30-m compensating reflector should be made". The cost of building a telescope of this type at a remote site was estimated to be greater than £11 million. In March 1970 the board accepted the committee's recommendation that certain aspects of the proposal required further research (which it agreed to finance). The 5 year forward look of the board prepared in 1970 for the years 1971–72 to 1975–76 referred to these discussions and the intention to award more grants for the further studies but stated explicitly that no provision has been included for the main facility.

Thus Dr Gebbie's recollections that the project was started or funded in that epoch are correct only in the above limited context. Furthermore, since the discovery of the 2.6 mm line radiation from galactic carbon monoxide was not made until 1970 (R.W. Wilson *et al. Astrophys. J.* **161**, L43–L44), this can have had no bearing on the issue of the UK facility at that stage.

The development of the proposals eventually leading to the present UK millimetre project belong to a later era and for the historical record the following summary may be of interest. In the spring of 1973 SRC asked Dr J.A. Saxton, the director of the Radio and Space Research Station (later the Appleton Laboratory) to prepare a case for a UK millimetre wave facility. Saxton presented his report to SRC in the summer of 1973. The main proposal was for a 12-m dish to be built by a German firm

for working the 2–3 mm waveband and to be sited at Sutherland in South Africa. By that time pressure was mounting for European collaboration on major scientific projects and SRC convened a "millimetre wave astronomy panel" with Professor A. Hewish as chairman to advise on the situation. When this panel met in January 1974 the general issue had been complicated because collaboration with Franco-German and Australian millimetre wave projects was also under discussion.

One problem facing the United Kingdom was the pressure on scientific manpower for any such extensions of its astronomical commitments and in the summer of 1974 R. St J. Walker, secretary of the Science Research Council asked me if we were able to make a substantial contribution to a millimetre wave programme from Jodrell Bank. Since this would have involved the diversion of staff from the development of the multi-telescope system (proposed after the cancellation of the Mk VA telescope) I answered with caution and reserve. However, in the early autumn of 1974 two other factors intervened. SRC agreed to support only the first phase of the multi-telescope system (MERLIN) and discussions with UK Atomic Energy Authority (UKAEA), which was acting as agent for the construction of our telescopes, indicated that at a relatively small cost the Mk II telescope at Jodrell Bank could be modified to a 100 foot circular aperture capable of operating down to a wavelength of 6 mm (the Mk IIA). On 9 October 1974 I despatched to SRC a formal proposal for this work and on the advice of UKAEA asked for £50,000 for the design study in 1975, and £1,500,000 for the construction in 1976–77. This request stimulated SRC to renewed activity and on 11 December 1974 all workers in millimetre wave astronomy were circulated with a request to make proposals for a UK millimetre wave programme.

In order to consider these various proposals, SRC convened a "working group to coordinate proposals on millimetre astronomy" with Professor M.J. Rees as chairman. The major result of the meetings of this group early in 1975 was to recommend that studies should be made for a UK facility capable of operating down to a wavelength of 1 mm. Funds of the order of £100,000 were made available by SRC for feasibility studies and in the autumn of 1975 SRC established a "UK national millimetre astronomy facility steering committee" under the chairmanship first of T.G. Phillips and later of A. Hewish. By the end of 1975, in the light of the feasibility studies, this steering committee considered that a 15-m aperture telescope to operate down to a wavelength of 0.75 mm would be feasible.

It was this sequence of events that led to the present proposal approved by SRC in

1980 for a 15-m aperture telescope with a working wavelength limit of 0.3 mm — that is a telescope one-half the size but working at a wavelength more than 20 times lower than that of the instrument referred to in the 1969 report of Professor Smith's working group.

BERNARD LOVELL

*University of Manchester,
Nuffield Radio Astronomy Laboratories,
Jodrell Bank, Macclesfield, UK*

SIR — H. A. Gebbie (*Nature* 24 March, p.286) is mistaken in suggesting a "failure to provide a rational basis" for siting the United Kingdom/Netherlands Millimetre-wave Telescope on Mauna Kea in Hawaii. The choice was made only after most careful consideration of all the relevant factors by representatives of the British and Dutch millimetre-wave astronomy groups. The evidence on atmospheric absorption was given very great weight in this analysis and it was the outstanding properties of Mauna Kea in this respect which led to its selection.

Gebbie refers to measurements made in 1976 by his group, which reported¹ strong excess absorption even in clear dry conditions. Since that time three large telescopes have come into operation on Mauna Kea and it has become established as the world's premier site for astronomical observations at wavelengths near 1 mm (refs 2–5). Measurements of the atmospheric transmission are made routinely as part of the calibration procedure for such observations. In the clear dry conditions which predominate at this site, the observers report^{6–13} good transmission and do not confirm the excess absorption claimed by Dr Gebbie's group. The general view is that, for submillimetre observations in particular, Mauna Kea is the best site in the world where a major observatory has been established.

To plan the scheduling of the telescope in a way which maximizes its efficiency, we need detailed measurements of the atmospheric transmission covering a long period. We have therefore developed an automated instrument for measuring the sky at wavelengths of around 1 mm. This was installed at the summit of Mauna Kea at the end of 1982 and the data collected so far confirm the excellent properties of the site.

RICHARD HILLS

*Cavendish Laboratory,
Cambridge, UK*

1. Moffatt, P., Bohlander, R.A., McCrea, W.R. & Gebbie, H.A. *Nature* **269**, 676 (1977).
2. Marscher, A.P. *Nature* **303**, 475 (1983).
3. Keene, J., Hildebrand, R.H. & Whitcomb, S.E. *Astrophys. J. Lett.* **252**, 1.11 (1982).
4. Phillips, et al. & Beichman, C.A. et al. in *Regions of Recent Star Formation* (Reidel, Dordrecht) (1982).
5. Lee, T.J. et al. *Nature* **295**, 214 (1982).
6. Cunningham, C.T. et al. *Icarus* **48**, 127 (1981).
7. Fetterman, H.R. et al. *Science* **211**, 580 (1981).
8. Lesurf, J.C.G. thesis, Univ. London (1981).
9. Padman, R., Hills, R.E., Cronin, N.J. & Rose, W.B. *Mon. Not. R. astr. Soc.* **192**, 87p (1980).
10. Riley, P.W. et al. *Mon. Not. R. astr. Soc.* **199**, 197 (1982).
11. White, G.J., Phillips, J.P. & Watt, G.D. *Mon. Not. R. astr. Soc.* **197**, 745 (1981).
12. van Vliet, A.H.F. et al. *Astr. Astrophys.* **101**, 1.1 (1981).
13. Robson, E.I. & White G.J. in *The Scientific Importance of Submillimetre Observations* (ESA, Noordwijk).

Nature conference

Thirty years of DNA

MOLECULAR biology seems, no doubt briefly, to be caught in the trap that gave early nineteenth century biology a bad name. The recognition that genes in higher organisms are interspersed with as yet apparently functionless stretches of DNA called introns seems to have stimulated a self-sustaining preoccupation with the most weird and the most wonderful. The piece most obviously missing from this jigsaw puzzle is a confident understanding of how genes became the kind of things they are, and how they control themselves (and each other).

This is one reaction to the conference to celebrate the 30th anniversary of the publication by J.D. Watson and F.H.C. Crick (*Nature* 171, 737; 1953) of a structure of DNA whose correctness has since been amply confirmed. The conference, organized by *Nature*, was held at Churchill College, Cambridge, on 13–14 April.

On such occasions, nevertheless, recollections of the past will out, advice to speakers notwithstanding. One confessed that he had been an intended mathematician when the Watson and Crick paper was published (W. Bodmer, Imperial Cancer

the problem of the human genome may be simply that of 1,000 to 3,000 gene clusters. The fact that actin-like genes have recently been found not always to specify proteins found in muscle seemed to Bodmer to be an illustration of how gene clusters would be defined by similarities of nucleotide sequences.

Multigene families cropped up repeatedly last week, most naturally in the summary by Lee Hood (Caltech) of what is known of the arrangement of genes in the 3 million base-pair region of mouse chromosome 17 which carries the genes for the type I and type II histocompatibility antigens as well as the genes encoding complement components, now known to be analogous to the HLA system in humans.

The rearrangement of the antibody genes is by now, of course, a familiar topic (see, for example, S. Tonegawa, *Nature* 302, 575–581; 1983). By supposing that the bits and pieces of the immunoglobulin genes on chromosomes 6, 12 and 16 in the mouse are rearranged in the lymphocytes that produce antibodies, and allowing for the occurrence of somatic mutation, it is possible to account for the capacity of the immune system to manufacture antibodies specific against an open-ended set of antigens: permutations available from among the several versions of each of several differently placed partial genes can easily explain 100,000 different antibody molecules even before allowing for mutation. The distinguishing feature of the family of antibody genes is their ability to undergo rearrangements which are mediated by conserved flanking sequences on either side of each of the pieces that eventually assemble to form a complete antibody gene. Could it be, asked Hood, that the similarities between these flanking sequences and those of "transposons" signal an evolutionary origin in migratory DNA, and the later exploitation of the migratory mechanism in the generation of antibody diversity? The source of the diversity of the transplantation antigens is, however, quite different, and consists of polymorphism.

Hood's purpose, last week, was to explain how this genetic diversity might have come about. In the mouse major histocompatibility complex (MHC) many more gene-like sequences have been detected than are needed to account for the known

Hood and Borst



Jim Watson, in his introduction to the conference, looked back . . .

WHY thirty years and not, say, twenty years or ten years? Well, it's pretty clear why it couldn't be ten years, because we waited 9½ years to get the Nobel Prize. Fifteen years wouldn't have had any meaning, as the big event occurred after thirteen years, in 1966, when the genetic code was solved. Then repressors came along, but that wasn't enough for a big celebration. And then twenty — we didn't know it at the time, but that was the real year, because that was when we came to recombinant DNA. It was made practical by Boyer and Cohen. People ask me "Why did the Swedes take nine years to give you the Nobel Prize if it was so important?" But they still haven't given it to Boyer and Cohen.

Nature did celebrate twenty-one years, and that was really very nice, because Francis [Crick] wrote and, in particular, Linus [Pauling] wrote. But we couldn't sell the twenty-fifth anniversary very big, because we were still mad at each other. We couldn't work on tumour viruses, because of the regulation of recombinant-DNA research. And it's really only now, in year thirty, that we're free to do exactly as we want.

Now, I guess, people like me have lots of debts to acknowledge — mostly to people that those here really didn't know. But there were two unique patrons — people who protect you if you're trying to do something different. One was the Rockefeller Foundation. It really started the thing. It gave money in 1933 to begin the first Cold Spring Harbor Symposium, and it gave a lot of money to Caltech and it gave money to the Medical Research Council (MRC). It gave money to Indiana University. It gave money to Copenhagen, when I was with Herman Kalckar. The Rockefeller Foundation did this because of one man, Warren Weaver, a mathematician from the University of Chicago, president of the University of Wisconsin, and . . . a



Bodmer and Williamson

Research Fund), another that he was anxious to be a chemical engineer (R. Williamson, St Mary's Hospital Medical School). Williamson was one of several European participants who drew attention to the time-lag in accepting the view that the structure proposed for DNA might be significant, quoting the then head of a British biochemistry department as saying that "this nucleic acid stuff will come to nothing".

Two session chairmen — Bodmer and Williamson — gave convincing reasons why molecular biology is, temporarily, natural history. Bodmer, once a student with R.A. Fisher, drew attention to the importance and antiquity of the concept of the gene cluster, defined by classical geneticists, but which is now quantifiable. In human genetics, the classical cluster is that of the globin genes, but now there are more telling systems such as those of the histocompatibility genes (HLA in human beings, H-2 in mice) and the immunoglobulin genes. If half of the 3×10^9 base pairs of the human genome consists either of "junk" or of selfish DNA, and if an average cluster contains about 15 genes,

fine man. It was his belief that helped Linus Pauling and George Beadle to talk about how chemistry was going to revolutionize biology. Weaver provided a spirit in which younger people could think.

The second patron was MRC. Harold Himsworth supported the MRC laboratory in its earliest days and Sidney Cattell, my better, saw that the new laboratory was built, without which Cambridge couldn't have been the outstanding place it still is today.

The second kind of patron is more personal. We had Sir Lawrence Bragg. As the Cavendish Professor, he could assign space. And he assigned some space to Max Perutz and then to John Kendrew so that there were six people in Cambridge when I arrived in 1951. And, as I recount in *The Double Helix*, I thought Bragg was just a stuffy old man when I met him. But he was a fine man. He had a really keen interest in science, and he was certainly Francis's only competition at that time — in the sense that he was a theoretician. And he had a difficult time, because most people thought that it was his father who had been the clever one, whereas it was the younger Bragg who'd made the running. When he came to the Cavendish, people said how dreadful it was that he was not a nuclear physicist or even a low-temperature physicist, but just a crystallographer. So his support of us was very important. I showed him several manuscripts, and he really was very helpful.

Then there was the patronage given to Francis and me by John Kendrew and Max Perutz. Francis was Max's research student and I was John Kendrew's. And when I was going to lose my fellowship for coming to Cambridge, they dug up enough money for me to stay in Cambridge.

But I guess I owe most of all to Francis, who really did look after me, and who often tried to keep me from being silly. I wasn't as silly as he thought, but he was so sensible that I had occasionally to say things I didn't believe, to see if I could trap him. And I sometimes did.

I don't think the whole thing would have worked if the Cricks hadn't cooked so many meals for me, or made me feel at home in Cambridge by seeing that I didn't cut my hair for quite a while. And then seeing that I wore a tie. And that I got an English suit. And giving me the good advice that I shouldn't look like an American. I followed their rules to the point where it made it difficult sometimes to go home.

I couldn't have got anywhere without Francis, so I really felt a little strange coming back here for this meeting, because it's without Francis. It could have been Crick without Watson, but certainly not Watson without Crick.

I wrote *The Double Helix* because it was a good story. Some people claimed they didn't recognize me in it, but others thought they recognized me too well. Originally, it had the title *Honest Jim*, but some people objected because they thought

polymorphic class I antigens and their non-polymorphic relatives, the lymphocyte differentiation antigens. This has given rise to the suggestion that the extra gene-like sequences in the MHC may represent a reservoir of genetic information that could contribute to the polymorphism.

Hood also advised looking for multigene families elsewhere. Will some gene cluster, or a few, explain the specificity of, for example, the detection of smell. Maybe. But can it be so simple?

This interest in movable genes, the malleable genome, is now more widely shared. The most striking illustration, last week, of the importance of the malleability of the genome was the account by Pro-



Hogness and Temin

fessor P. Borst (Amsterdam) of the mechanism of antigenic variation by trypanosomes, the bloodstream parasites responsible for sleeping sickness. In the course of a trypanosome infection, as many as a hundred changes may be rung on the antigen specificity of the protein component of the outer envelope. Hybridization studies involving cDNA and native DNA are consistent with the view that trypanosomes carry a stock of silent genes which spring into action only when they migrate to the neighbourhood of what Borst calls "an expression site".

How is this accomplished? As Borst explained, it appears that there is more than one way of getting a new antigen gene into the "expression site", which is right at one end of a trypanosome chromosome. One possible mechanism is reciprocal exchange with the end of another chromosome, replacing the expressed antigen gene but leaving the rest of the chromosome intact. To have the requisite number of silent genes at these ends of chromosomes, trypanosomes carry several hundred "mini-chromosomes".

As it happens, trypanosomes have only six times as much DNA as *Escherichia coli*, much of which must be given over to the genetic ingredients of antigenic variability (but trypanosomes cannot live by agar alone). *Drosophila melanogaster* is, by contrast, more like the human genome in its complexity, with the further attraction



of carrying known genes which regulate the development of the organism by some means still unknown. David Hogness described the success with which his colleagues at Stanford have nevertheless been able to clone the huge 75,000 base-pair region of the *Drosophila* genome in which homeotic mutations of the bithorax complex of the *Ultrabithorax* class are known to be scattered. The surprise is that most of the spontaneously arising mutations appear to entail the insertion of transposable elements into two short segments towards the ends of the 75,000 base-pair transcriptional unit that apparently encode the major exons of the discrete RNA products of this gene. The sheer size of this transcription unit remains something of a puzzle, and Hogness wondered whether it could be a mechanism for timing developmental events.

Movable genetic elements were also conspicuous parts of last week's account by Howard Temin (University of Wisconsin) of the molecular biology of the retroviruses capable of transforming cells into a malignant condition. That RNA can carry genetic information is well established. Temin asked his audience to consider exogenous retroviruses as being in a kind of equilibrium with the corresponding genetic information (written in DNA) in the genome of the infected cell, which is in turn potentially a movable element.

Temin's own work with the only retrovirus known to infect turkeys has persuaded him of the complexity of sequence of genetic transformations that must precede the conversion of cellular genetic information into the viral oncogene, involving the fashioning of a new promoter and the beginning of the gene, the elimination of intervening sequences and a variety of point mutations. Perhaps because of these changes, when the viral oncogene is integrated into the cellular genome, it is at a chromosomal site different from that of the progenitor gene. The problem of the genetics of cancer, he said, is far from understood.

R. Jaenisch (Hamburg) has used retroviruses for a different purpose, to discover into which embryonic genes they are inserted in such a way as to create a lethal, recessive mutation. By breeding experiments in which hundreds of mice have been involved, Jaenisch has so far discovered one such insertion which, in homozygous mice, produces intrauterine death at about day fifteen.

Why should that be? Jaenisch appears to have had the good fortune that the viral gene in the affected animals had been inserted in a region of the mouse genome coding for one type of collagen. As it happens, the viral gene is incorporated backwards, at the beginning of the collagen gene, which may be one reason why homozygous embryos fail to produce the protein. But Jaenisch says that this cannot by itself be a complete explanation.

In the natural history of genes, catalogu-

ing occurrences along linear stretches of DNA is the obvious first step. Telling how the DNA is disposed in chromosomes has turned out to be a harder task, but the conference last week was told (by Aaron Klug, MRC Laboratory of Molecular Biology, Cambridge) of new X-ray diffraction data from single crystals of nucleosomes — units of chromatin consisting of 140 base pairs of DNA wound in $1\frac{3}{4}$ turns of a left-handed superhelix around a core of eight protein histone molecules.

The Cambridge group has obtained X-ray diffraction data at a resolution of 7 Å, at which point the disposition of the two grooves (major and minor) of the helix can be identified. The windings of the molecule on the histone core are closer together than expected. The problems of understanding how such compact structures are replicated and transcribed remain.

Another way of understanding how DNA molecules dispose themselves in real life, described last week by J. Wang (Harvard), smacks more of physical chemistry. Take a circular molecule of double-stranded DNA from polyoma virus, and artificially insert bits and pieces of DNA so as to tell (from the mobility of the molecules thus formed) how the degree of supercoiling has been affected.

Wang described how stretches of the repeated C-G (cytosine, guanine) sequence do indeed affect the mobility of the circular molecules in a way that suggests that a piece of Z-DNA (left-handed) has been formed. He has also been able to calculate the free energy of formation (60 kcal per mole) of cruciform structures formed by the insertion of an artificial sequence of 68 base-pairs. Wang's plan now is to study the physiological significance of supercoiling with an artificial plasmid containing the gene for topoisomerase from *E. coli*.

Mere description, unfortunately, does not explain how genes may actually function. The best known piece of naturally occurring genetic machinery remains the working of the lambda repressor, the gene carried by the lambda-phage which, in *E. coli*, can turn off the genes responsible for replication of the virus. The principal present custodian of this piece of machinery, M. Ptashne (Harvard) described how the N-terminal ends of pairs of repressor molecules attach themselves to one of the three contiguous sites in the major groove of a DNA helix so that part of the attaching protein embraces the whole helix.

By model-building and with the help of lambda mutants whose structural consequences can be predicted, a virtually complete picture of the stereochemistry of repressions has been built up. Ptashne says that his immediate ambitions include making mutants of the bacterial RNA polymerase to confirm that repressor molecules control its activity in the way now predicted, to tinker with the binding sites of the repressor molecules and to study newly obtained co-crystals of the repressor of bacteriophage 434 and

bacterial DNA.

The functions (if such there are) of intervening sequences remain unclear. How, in any case, did they arise? C. Weissmann (Zurich) gave it as his opinion that there is no way of telling whether genes with introns are more or less "primitive" than those without. Virus and bacteria lack them, but they survive.

The stability of introns within particular genes argues for function (although the frequent appearance of particular sequences within introns is more probably necessary for the accuracy of splicing). Weissmann took the view that introns are very old, perhaps parts of the earliest RNA genomes, and that they must during the



Ptashne and Wang

evolution of different genomes have been a valuable means of amplifying functional genes and of making possible genetic rearrangements without breaking up coding sequences of DNA. How else could the family of immunoglobulin genes have arisen? But with all this said, the complexity of the collagen gene remains a puzzle. Perhaps (as Weissmann said in another context) "Nature is trying to tell us something".

W. Gilbert (Biogen) put forward last week another way of coming to grips with the same question — the notion that the functional parts of split-up genes have indeed been recruited from elsewhere, and that as a consequence the distribution of introns should be reflected in the structure of the proteins eventually produced.

The notion that protein molecules consist of recognizably distinct domains is not of course new. Gilbert, however, was able last week to show that the view has some predictive value, in spite of some disconcerting evidence — the recognition, for example, that the eight-fold symmetry of the enzyme triose phosphate isomerase (from the chicken) is coded for by seven and not eight gene pieces.

To describe all this as natural history is, of course, unfair. For the phrase applies only in the sense that while genetic maps are still incomplete, most new information starts fresh hares. And there have already been several telling illustrations of how it is



Weissmann and Gilbert

I was proclaiming myself as the Honest Truth. So then I changed it to *Base Pairs*, and that didn't go over either, so the book ended up with the title *The Double Helix*. I had this great idea of a picture of me looking at Francis and Maurice [Wilkins] looking at Rosalind [Franklin] on the cover of *Base Pairs* — a kind of *Kind Hearts and Coronets* arrangement. But it ended up a little more sensible.

Rosalind Franklin was a very intelligent woman, but she really had no reason for believing that DNA was particularly important. She was trained in physical chemistry. I don't think she'd ever spent any length of time with people who thought DNA was important. And she certainly didn't talk to Maurice [Wilkins] or to John Randall, then the professor at Kings. And then the time came when he moved her out of DNA and sent her over to Bernal, and she had to write up the papers.

As Aaron Klug has explained in *Nature*, she came very close to the structure of DNA. She really had accepted that it was a helix, but we didn't know that. She was moving towards the idea that there were two chains, but she never built models. Now if she could have had just two hours with Francis, and she could have been convinced that he was right, not just a loud-mouth, I think she would have gone back to her laboratory and built models. She'd have solved the structure of DNA. But she was really prepared to give up working on DNA, and she wouldn't have agreed to give up if she'd thought it was important. So that was why she didn't get the answer.

Probably none of you knows how much she liked Francis. Afterwards she came to him to talk. She did that very beautiful work on tobacco mosaic virus (TMV) and when she had her operation for ovarian cancer, which she knew was very serious, she came and stayed with the Cricks to recuperate. She was pretty ill, but she was supported strongly by Don Caspar, and less so by me, and we got her a National Institutes of Health grant that let her do the TMV work when the Agricultural Research Council (ARC) turned her down after she'd told the head of ARC that he was an idiot. He was an idiot. When she wanted a diffractometer, he pointed out that there was one in Aberdeen. So she told him what an idiot he was and then they didn't support her. She wasn't a diplomatic person but when you got to know her she was fine to deal with.

Linus [Pauling] didn't deserve to get the structure. He really didn't read the literature. And he didn't talk to anyone either. He'd even forgotten his own paper with Max Delbrück which said that a gene should replicate by complementarity. He seems to consider that he should have got the structure because he was so bright, but really he didn't deserve it.

Now he might have got it because Alex Rich arrived at Caltech and began to take X-ray photographs at just about the time that we proposed our model. I think it was

inevitable that the structure would have been solved within about a year. The momentum was there, and they knew that DNA was important.

But I have no guilt feelings. Both Francis and I were products of a tradition that wanted to solve the problem, right through. I'd been trained in the phage group, for which self-replication was the key problem. And Francis was the heir to the tradition established here in England that molecular structure is important, and if you work out bigger and bigger structures, you'll learn something important.

What happened when we got the structure? Well, we wrote the paper for *Nature* and then we realized that the final throw-away line might have been too cute and that we'd better write a longer paper. So, so that some third person wouldn't write down these trivial ideas and claim credit for them, we wrote the second *Nature* paper called "The genetical applications of the structure of DNA". And then Francis started to talk. I mean he'd always talked. And then Sydney [Brenner] came over. On about the first occasion I saw Sydney, we talked for about six hours non-stop. Sydney had a few bright friends in Oxford who talked about DNA, but they didn't have expert pictures and didn't do anything about it. Sydney was then a research student with Hinshelwood, who was a kind of Lamarckian and . . . a strange man. And Sydney was doing a PhD thesis that was as dull as Francis's, on bacteriophage T4. But he knew the phage was important.

He took me aside one day and said "Jim, you don't realize how important the work you've done is". I think I did, but I was also scared — it might not be right. One had this sort of feeling because of the Cambridge biochemists who were calling it the WC structure. When Francis got a request to speak on BBC radio, I said that he could do it if the broadcast went out on the External Services and if it wasn't heard in England. So his voice talking about DNA and its importance was heard only outside England. But then I realized that it was the end of an era. I'd been raised at courses on H.J. Muller, and I'd heard what genes might be and how they might self-replicate. So when we saw the complementary structure of DNA, that was the final solution. We didn't see it, as we now see it, as the beginning.

Both Francis and I had no doubts that DNA was the gene. But most people did. And again, you might say, "Why didn't Avery get the Nobel Prize?" Because most people didn't take him seriously. Because you could always argue that his observations were limited to bacteria, or that [the transformation of *Pneumococcus* that he described was caused by] a protein resistant to proteases and that the DNA was just scaffolding. Only when they saw the double helix did people stop asking what the gene was.

The first thing I worried about was whether the strands would come apart

quickly enough and how to explain [Arthur] Kornberg's discovery that the chains run in opposite directions. But I doubt whether Francis and I combined spent more than ten hours worrying whether the structure was right, whether they would take it away from us or something like that. So it was very satisfying when we saw that the replication scheme seemed to be right, even though [L.F.] Cavalieri came along and said we should use a four-chain model.

What happened next? There was the area of the genetic code, or how information was included in DNA. That was dominated by Francis and Sydney, Khorana and a few others, but I'd say the work here at Cambridge was the intellectual high-point of that whole thing. Then we were all dominated by the biochemistry of the central dogma [DNA makes RNA makes protein] and we all wanted to find the enzymes that did all those steps. That was an era when we were the new enzymologists. But when Francis and I saw the structure, we'd never thought in terms of an enzyme. We even thought that because it's a template, you might not need those enzymes. But now, of course, there's this great RNA splicing story without enzymes that makes one think that we weren't totally mad. But it was far out, and the person who used to argue with us most about enzymes was Peter Mitchell, which was interesting because [in his later work] Peter managed to do away with the need for several enzymes.

Without first class biochemistry of this sort, from first class conventional biochemists such as Lippman and Ochoa, we wouldn't be here today, because the enzymes were necessary. But then we also needed the wonderful molecular genetics which came from people such as Benzer, Sydney [Brenner], Jacob and Monod and which finally led to Mark [Ptashne] and Wally [Gilbert] and the repressor story.

But then we didn't know what to do next. Many of us moved into tumour viruses, hoping that what had worked for phage and bacteria would work for tumour viruses in animal cells. It's chilling to think of where we'd be if the cloning of DNA hadn't come along. We would be stuck so far back that it wouldn't be much fun. Even there we owe an enormous amount to a genetics approach that has really made us think about plasmids and so on until, finally, recombinant DNA came along.

But then we were involved with the question of whether we had something so powerful that we shouldn't do it. We'd gone through that. And [the fuss about the control of recombinant-DNA research] is now so far back that I don't really feel mad. I always regarded the regulation of recombinant-DNA research as a black comedy, and said so. But some people thought the concern was genuine. It was genuine for some, but for others it was really the theatre of the absurd. And it's over. What a relief!



Nasmyth and Jackson

possible deliberately to explore the genome for particular functions.

Much in this spirit, J. Jackson (Columbia) is one of the growing band embarking on the exploration of the nervous system. He described work in which he and his associates have identified and cloned genes encoding three previously identified neuropeptides that control the behaviour at egg-laying of the marine invertebrate *Aplysia*. These genes generate polypeptides that can be enzymatically processed to yield not only the three isolated peptides but, potentially, a host of others.

Gene regulation is another preoccupation. Last week, K. Nasmyth (MRC, Cambridge) described his hunt for the means by which the switch from one mating type to another is accomplished in yeast. H. Pelham from the same laboratory demonstrated the power of the techniques now available for incorporating genes into animal cells (HeLa cells and *Xenopus* oocytes, for example) as a means of identifying control elements upstream of the point at which transcription begins.

Biotechnology apart, Y.W. Kan (San Francisco) gave an account last week of how molecular biology can prevent genetic disease, by the use of genetic probes to detect patterns of inheritance.

So where will it lead? Sydney Brenner (MRC, Cambridge) offered a tantalizing vision of what may lie ahead. There may be Pelham, Brenner and Kan



no commanding heights still to be scaled, but only an apparently endless succession of foothills. Perhaps, "when we've done thymidine kinase and all the other enzymes we can think of", there will be little left.

So is there a magic room in which the secrets of life will be found? "We open one door, and find ourselves in a long corridor, at the end of which there is another door, and then another corridor, and so on." That, Brenner said, is what science is like. The next breakthrough? The question may be too late. Perhaps the breakthroughs have already been made and we need not await another Galileo, "the Watson and Crick of development", Brenner said. "Maybe that's what it's really like. People will just publish more and more . . ." □

Geophysics

Plate tectonics on Venus?

from R.J. Phillips

OVER the past few years, a wealth of radar altimetry data has been collected about Venus from the Pioneer orbiter and there has been much lively speculation as to whether the tectonic processes that operate there are similar to those seen on Earth. A valuable review of the present situation has recently been provided in the *Journal of Geophysical Research*¹ by Sean Solomon of MIT and James Head of Brown University; their conclusion is that it is still not possible to rule out any of the types of tectonics that might plausibly be operating on Venus.

In general, planets of the same size and mass — as are Earth and Venus — can be expected to have similar histories of heat flux from the interior. As the tectonics of terrestrial (predominantly silicate) planets merely reflects the style in which the outer rigid lithosphere is able to expel the planet's internal heat, then it is a reasonable first prediction that the tectonics of Venus will be broadly similar to those of Earth and be dominated by plate tectonics — seafloor spreading and continental drift.

Comparison of Earth with Venus is particularly valuable because it offers an opportunity to sort out the effects of the variables that take on different values in the two planets. We know, of course, that terrestrial plate tectonics is controlled by more than just the quantity of heat coming out of the Earth's interior, but just how important the various second-order geophysical variables are is a source of considerable disagreement.

To compare Venus and Earth, some assumption must be made regarding the relative concentrations of the heat-producing elements — potassium, uranium and thorium. The most plausible hypothesis — based on both cosmochemical arguments and elemental measurements made by the Soviet Venera landers² — is that the heat production per unit mass is approximately the same for both planets. This assumption, inherent in the rule that planets of equal size have the same thermal history, provides a basis for comparison of tectonic styles.

Solomon and Head propose that lithospheric heat transfer in the terrestrial planets can be represented on a pseudoternary diagram whose vertices are lithospheric conduction, plate recycling and hot-spot volcanism. Lithospheric conduction is heat transport by diffusion through the lithospheric shell; Mars, Mercury and the Moon lose their heat in this way. Plate recycling is heat loss by the creation and subsequent cooling of new hot lithosphere formed at divergent plate boundaries; plate consumption at convergent boundaries completes the recycling process. This is the

dominant mechanism of heat loss in the Earth. Hot-spot volcanism is the loss of heat by the direct access of magmas to the surface by volcanism; the subsequent cooling of these magmas provides the heat loss. The Galilean satellite Io appears to lose most of its heat by this means.

Presumably, some combination of these mechanisms provides the heat loss for Venus. Arguments against the presence of any plate tectonics on Venus, however, have included (1) that topographical features indicative of plate tectonics have not been observed on Venus; (2) that most of the surface of Venus is very old — the opposite of what would be expected with plate recycling; (3) that the high surface temperature makes the venusian lithosphere more buoyant and less likely to sink (that is, subduct) back into the mantle; and (4) that because of the high surface temperature, plate recycling would be less efficient for removing heat on Venus than on Earth (implying that Venus would then 'choose' a more efficient mechanism).

Solomon and Head claim that none of these four arguments precludes plate tectonics. The first argument, the failure to identify plate tectonic topographical features on Venus, has been the subject of considerable debate^{3,4} based on what could be seen given the topographical resolution of the Pioneer Venus altimeters after correcting for environmental differences at the surface (higher surface temperature, lack of an ocean on Venus). There seems to be agreement, however, that at least some mid-ocean ridges (the divergent plate boundaries of seafloor spreading) ought to be seen, although it has been argued⁵ that the more subdued of these would not be identifiable, given the errors in the altimetric measurements. The issue boils down to this: all seafloor spreading forms can be divided into one of two classes — those that can be resolved with the Pioneer Venus altimetry and those that cannot. If no such features can be identified in the altimetry, then plate tectonics cannot be precluded by merely assigning all plate tectonic features to the second class.

Earth-based radar images do, however, reveal features that are reminiscent of continental plate tectonic features on Earth. The best analogy is seen in the mountain ranges of the Ishtar Terra region, where radar images show bands with individual widths of 10–20 km. The banding is indicative of horizontal stress in the venusian lithosphere, which on Earth is associated with plate tectonics.

R.J. Phillips is in the Department of Geological Sciences, Southern Methodist University, Dallas, Texas 75275 and the Lunar Planetary Institute, Houston, Texas 77058.

Evidence that the lithosphere on Venus has not been recycled because the surface is old comes from circular and quasi-circular surface features which seem to have been caused by impact. Given our understanding of the impact flux history of the inner Solar System, the surface density distribution of these circular features implies that much of the surface of Venus is geologically old. This argument may be simply countered by noting that the circular features could just as well be of tectonic origin.

The buoyancy argument — that lithosphere is more buoyant on Venus — was never meant to rule out plate tectonics on Venus, only to point out that the lithosphere was less likely to subduct⁵. The observation made by Solomon and Head that subduction on Earth is essentially independent of buoyancy (that is, lithosphere age) begs the issue of buoyancy requirements for the initiation of subduction. As they correctly point out, young buoyant lithosphere on Earth that is subducting may be attached to older less buoyant lithosphere (which may have been involved in the initiation phase).

Finally, the efficiency of seafloor spreading as a means of removing the internal heat of Venus has been questioned, since the tectonics that actually occur presumably reflect the easiest way to get heat out of the interior. If the major topographical features of the equatorial regions — Beta Regio and Aphrodite Terra — are assumed to be spreading ridges, then one of two possibilities exists. If free subduction takes place, then thermal boundary layer theory predicts that the ridges decrease in elevation much too rapidly away from their axes, assuming equivalent heat source concentrations for Venus and Earth. Alternatively, heat sources on Venus may be extremely depleted. A second possibility is that the sharp decrease in elevation is due to a spreading velocity that is inhibited relative to that predicted by theory. Such is the case on Earth for lithospheric plates which in part include continents. If subduction is inhibited on Venus, then the slower velocities and lower total 'ridge' length (as in the Equatorial Highlands) suggest that seafloor spreading could deliver only 15 per cent of the global heat flux⁶, as compared with 65 per cent on Earth.

Solomon and Head suggest that seafloor spreading, if it exists, is probably confined to the vast plains of Venus. The spread of elevations in the plains, however, is too great to be explained by cooling of the thermal boundary layer of seafloor spreading⁷, so if seafloor spreading exists in the plains, it is subtle and dominated physiographically by other tectonic processes — perhaps volcanism and/or hot-spot swells.

Of the three tectonic end members defined by Solomon and Head, it can be agreed that none can be ruled out. In fact, at least two can be ruled in. That is, it would be very hard to avoid conductive heat loss through the lithosphere. Further, radar

interpretation of shield volcanoes and Soviet Venera measurements of surface elemental abundances matching terrestrial basalts both point to volcanism and the hot-spot mode of heat loss.

Of the oceanic and continental aspects of plate tectonics, only physiographical evidence consistent with the latter exists — in Ishtar Terra. The former, if it exists, would seem to be confined to the plains and to be topographically subtle. An imaging radar mission is required to resolve the

issue; fortunately such a mission appears in the 1984 budget proposed by the Reagan administration. □

1. Solomon, S. & Head, J. *J. geophys. Res.* **87**, 9236 (1982).
2. Surkov, Y.A. *et al. J. geophys. Res.* **88**, (1983).
3. Arvidson, R.E. & Davies, G.F. *Geophys. Res. Lett.* **8**, 741 (1981).
4. Head, J.W., Yuter, S.E. & Solomon, S.C. *Am. Sci.* **69**, 614 (1981).
5. Phillips, R.J. & Malin, M.C. in *Venus* (eds Hunten, D.M. *et al.*) (University of Arizona Press, in the press).
6. Kaula, W.M. & Phillips, R.J. *Geophys. Res. Lett.* **8**, 1187 (1981).
7. Morgan, P. & Phillips, R.J. *J. geophys. Res.* (in the press).

Oncogenes

We still don't understand cancer

from Howard M. Temin

By now only a very uninterested reader of *Nature* could have failed to note the excitement about oncogenes and retroviruses (RNA tumour viruses). Oncogenes are pieces of modified cell DNA found in certain highly oncogenic retroviruses. Proto-oncogenes are the DNA sequences in normal cells related to oncogenes. The transcripts of proto-oncogenes are present in different tissues at different times in development¹. A popular current hypothesis in cancer research is that activation of proto-oncogenes results in cancer. This hypothesis is applied to cancers caused by strongly and weakly oncogenic retroviruses and to cancers with other causes. Experiments to test this hypothesis and see whether it can be extended to human cancers are underway in many laboratories.

The hypothesis was supported by studies of lymphomas induced in chickens by weakly oncogenic avian retroviruses, that is, retroviruses without an oncogene²⁻⁴. There it was found that in over 85 per cent of the tumours, avian retrovirus DNA was integrated near a proto-oncogene, *c-myc* (the proto-oncogene related to the oncogene of avian myelocytomatosis virus), and the proto-oncogene mRNA was present at higher concentrations than in normal cells.

One paper which has recently appeared in *Nature* presents further data which do not fit the proto-oncogene activation hypothesis quite so neatly. Tschilis *et al.*⁵ show that lymphomas induced in rats by weakly oncogenic murine retroviruses do not fall simply into the pattern predicted by this hypothesis. They do, however, find that the integration sites for murine leukaemia viruses in virus-induced lymphomas are not completely random; five of sixteen tumours had proviruses in the same region of DNA.

In another experiment to look for alteration of proto-oncogenes, mouse mammary tumour virus, another slowly oncogenic

retrovirus, was found frequently integrated at the same locus in the genome of mammary tumours, although no proto-oncogene has yet been reported in this region of DNA⁶. Thus Tschilis *et al.*⁵, as well as Yoshimura and Levine⁷, looked for integration of virus DNA in a single locus in murine leukaemia virus-induced tumours. However, it can only be concluded from their results that there is no single major locus like *c-myc* for the site of murine leukaemia virus integration in murine virus-induced lymphomas.

The concentration on sites of retroviral integration in tumour cells has left the problem of retroviral integration in normal cells behind. Numerous studies of both proviruses from large populations of cells and small sets of molecularly cloned proviruses support the hypothesis of random integration of retrovirus DNA. However, Cohen and Murphey-Corb, in another recent paper⁸, have challenged this conclusion for baboon endogenous retrovirus in human cells. They report a major band of viral DNA after *Pst*I digestion of DNA from chronically infected human cells. No major band was found after digestion with other restriction enzymes. They propose that integration in one orientation in a small region of repeated DNA containing a *Pst*I cleavage site is preferred for baboon endogenous virus DNA integration in human cells. However, their results might be more relevant to the study of sites of integration of proviruses in tumour DNA. In both cases cells were studied long after the initial infection, so selection rather than 'targeted integration' could be the basis for the apparent specificity of integration.

Also relevant could be the still unexplained phenomenon reported with spleen necrosis virus, an avian retrovirus, that *Eco*RI digestion of DNA from chronically infected, but not acutely infected, avian cells which does not cut viral DNA, results in different sizes for infectious and noninfectious DNA molecules⁹. Thus, these studies of retrovirus integration in chronically infected cells may only indicate that there is more order in the cell genome than is appreciated.

In any case the activation of a proto-oncogene by a tumour virus is but one of six ways in which proto-oncogenes have been shown to be genetically altered. The complete list is (1) by spontaneous mutation and insertion between the controlling sequences of a retrovirus which results in a highly oncogenic retrovirus with an active oncogene, (2) by insertion of a retrovirus adjacent to a proto-oncogene, (3) by attachment in a test tube of a large terminal repeat to a proto-oncogene, (4) by a spontaneous base-pair mutation, (5) by gene amplification and (6) by spontaneous translocation to a region of another chromosome which may be transcriptionally active. Only proto-oncogenes altered by mode (1) have been shown to cause neoplastic transformation of normal cells. Some of the proto-oncogenes altered by modes (3) or (4) can cause transformation of a preneoplastic line of mouse cells (NIH 3T3 cells).

c-myc has been shown to be altered by modes (2), (5) and (6). These alterations may result in increased transcription of *c-myc*, but they do not result in the ability of the altered *c-myc* to cause transformation of NIH 3T3 cells. In several instances proto-oncogenes activated by modes (2) and (6) do not cause transformation of NIH 3T3 cells while other DNA sequences from the same cells do^{10,11}.

These findings are consistent with the idea that events additional to the alterations of proto-oncogenes are required for the development of neoplasia.

Two possible examples of these additional events are the deletion of suppressor sequences required for the expression of the viral oncogene in reticuloendotheliosis virus strain T¹² and the homozygous deletion in human chromosome 13 apparently required for retinoblastoma¹³.

Thus, in spite of the dramatic findings of activation of proto-oncogenes, transforming DNA and viruses associated with human leukaemia, cancer remains what the classical oncologists have always believed it to be — the final product of a multi-step process. The still unanswered questions are: what are these steps, which of them are limiting and which are most susceptible to intervention?

One can only conclude that all the answers are not yet in and hope that the next years will be as full of illuminating discoveries as the past ones. □

Howard M. Temin is in the McArdle Laboratory for Cancer Research, University of Wisconsin, Madison, Wisconsin 53706.

1. Muller, R. *et al. Nature* **299**, 640 (1982).
2. Hayward, W.S. *et al. Nature* **290**, 475 (1981).
3. Payne, G.S. *et al. Nature* **295**, 209 (1982).
4. Noori-Daloui, M.R. *et al. Nature* **294**, 574 (1981).
5. Tschilis, P.N. *et al. Nature* **302**, 445 (1983).
6. Nusse, R. & Varmus, H.E. *Cell* **31**, 99 (1982).
7. Yoshimura, F.K. & Levine, K.L. *J. Virol.* **45**, 576 (1983).
8. Cohen, J.C. & Murphey-Corb, M. *Nature* **301**, 129 (1983).
9. Battula, N. & Temin, H.M. *Cell* **13**, 387 (1978).
10. Cooper, G.M. & Neiman, P. *Nature* **292**, 857 (1981).
11. Crews, S. *et al. Science* **218**, 1319 (1982).
12. Chen, I.S.Y. & Temin, H.M. *Cell* **31**, 111 (1982).
13. Benedict, W.F. *et al. Science* **219**, 973 (1983).

Microbiology

Antibiotic resistance and the evolution of bacteria

from Mark Richmond

A PAPER by Victoria Hughes and Naomi Datta in this week's issue of *Nature* (p.725) fills a gap in our knowledge of the emergence of antibiotic resistance in bacteria. The study makes use of a remarkable collection of bacteria made between 1917 and 1954 — before the use of antibiotics became widespread and at a time when little was known of bacterial genetics — and kept in sealed vessels since that period.

The study of bacterial plasmids and episomes, vital to an understanding of the spread of antibiotic resistance, did not really begin until the 1960s. Early on, most experimental work was focused on a few bacterial species — notably *Escherichia coli* and *Staphylococcus aureus* — and the plasmids/episomes involved were limited to a few types of phage and antibiotic-resistance factors and to some colicin factors. What emerged was a picture of the bacterial genome in which most of the genetic information in the organism was seen to be concentrated as a single large circular DNA molecule — the chromosome — of molecular weight about 2×10^9 , but some was carried as small circular DNA molecules which did not seem absolutely necessary for bacterial growth and survival in most environmental circumstances. When these extrachromosomal DNA elements carried the relevant genetic information they behaved as bacterial viruses, capable of being triggered to multiply rapidly to burst the host cell.

A further characteristic of many plasmids and phage was their ability to transfer from one bacterial strain or species to another and, in some cases, to catalyse the transfer of further independent circular DNA molecules. At the extreme, certain bacterial plasmids, following integration into the bacterial chromosome, could even cause the transfer of a copy of a whole chromosome from one bacterial cell to another to give rise, at least transiently, to a diploid organism.

Our awareness of the range of possible interactions of bacterial plasmids and the effects these interactions had on the bacteria radically altered our view of the individuality of single bacterial cells. Populations of bacteria could now be seen as a single gene pool, with most of the information segregated for much of the time in the individual members of the population, but with extra DNA molecules distributed

within the population and capable of being transferred within the members of the population. Furthermore, that integration and excision of plasmids into and from the chromosome sometimes transferred blocks of information from the chromosomal to the extrachromosomal state underlined the fact that a bacterial population as a whole had considerably genetic fluidity even if, under normal circumstances, this fluidity was constrained. When it was realized that this potential fluidity extended to populations containing different species and genera it was clear that one of the main processes determining the means of bacterial evolution had been discovered and that a much deeper understanding of the emergence of antibiotic resistance could be gained.

The emergence of resistance had already been broadly defined: the introduction of a novel antibacterial agent was followed by the rapid appearance of resistant bacterial populations which then became more and more widespread, both geographically and also with respect to the bacterial strains and

species exhibiting resistance, as the use of the agent continued.

It became clear that two processes are involved in the evolution of resistant populations: the 'invention' of the resistance genes themselves, and their multiplication and spread. Relatively little is known even now about the first of the two stages, though there are some clues. As for the second, the mechanisms of genetic exchange discussed above, coupled with selection, can provide a complete explanation.

One part of this story has, however, always been taken for granted, and it is here that the paper appearing in this week's issue of *Nature* finally provides definitive evidence. For the scenario outlined above to be correct, bacterial strains isolated before the advent of large-scale use of antibiotics should exhibit all the genetic flexibility of the antibiotic-resistant bacteria we see nowadays, but the incidence of resistance genes themselves should be low. Using a large collection of Gram-negative bacterial isolates made by R.G.E. Murray from before the antibiotic era, Naomi Datta and Victoria Hughes have now shown that this is indeed the case.

In a sense, therefore, Murray's collection is representative of the kinds of bacterial population that existed when antibiotic therapy started in the early 1940s. Plasmid transfer and gene exchange supported by appropriate pressure have done the rest. □

Human genetics

Is there a human *T/t* locus?

from P.N. Goodfellow and P.W. Andrews

CHROMOSOME 17 of wild mice frequently contains a region of mutant genes, known as the *T/t* complex, maintained in association by local suppression of recombination. Within the region, which stretches from the centromere to beyond the *H-2* locus, different *t*-haplotypes are characterized by different *H-2* types and different recessive lethal mutations with striking effects on embryogenesis¹. The *t*-haplotypes also carry a group of interesting mutations which, when homozygous, cause male sterility and when heterozygous cause carrier males to transmit the chromosome carrying the *t*-haplotype to nearly 100 per cent of their offspring. A similar constellation of mutations known as *Segregation Distorter* has been observed in wild populations of *Drosophila melanogaster*. Evidence for a locus homologous to the *T/t* locus in mammals other than mice, and perhaps rats², has been harder to find. Several recent papers suggest that a human homologue of the *T/t* complex may indeed exist³⁻⁶.

A human equivalent of the murine *T/t* complex would be expected to be associated with segregation distortion, recombination suppression and recessive

lethality. It might also be expected that an equivalent of the *T/t* complex would be linked to the human equivalent of *H-2*, *HLA*. This possibility has provided the starting point for several searches for the human *T/t* complex. Genes known to be linked to the *HLA* region include loci controlling class 1 and class 2 cell-surface antigens, loci controlling complement components and a locus controlling the enzyme glyoxalase (*GLO*). At least two alleles are known for each of the four major histocompatibility complex (MHC)-linked complement loci, but these are so closely linked that recombination between them has not been observed and only 14 common *cis* combinations of alleles (complotypes) occur.

In a survey of families drawn from the Boston area, Awdeh *et al.*³ observed that several complotypes showed non-random association with alleles at the *HLA-D* and *HLA-B* loci and less often with alleles of

P. N. Goodfellow is in the Laboratory of Human Molecular Genetics, Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX and P. W. Andrews in the Wistar Institute, 36th Street and Spruce, Philadelphia, Pennsylvania 19104.

Mark Richmond is Vice-Chancellor and Professor of Molecular Microbiology in the University of Manchester, Manchester M13 9PL.

the *HLA-A* locus. These associations have been called 'extended haplotypes', and the suggestion was made that segregation distortion and cross-over suppression comparable with the *T/t* complex could be the underlying cause. However, linkage disequilibrium between the *HLA* loci *A*, *B* and *D* has been known and studied for many years, and over the short genetic distance involved (2 centimorgans) it could be caused by many other mechanisms (including selection, recent mutations, founder effects and non-random matings)⁷.

More compelling evidence of a human *T/t* complex is presented in the second part of the paper. Combining data from several families segregating for one of the extended haplotypes (denoted *B8*, *DR3*, *SCOI*), it was observed that the haplotype was transmitted by males to their offspring significantly more often than the expected Mendelian property of 50 per cent, whereas transmission from females was not significantly different from 50 per cent. This was the case only when the *B8*, *DR3*, *SCOI* haplotype was in *cis* association with the *GLO-2* allele (5 centimorgans proximal); normal segregation was observed when the extended haplotype was associated with the *GLO-1* allele. The suggestion was made that in this population, the *B8*, *DR3*, *SCOI*, *GLO-2* haplotype contains a linked factor causing male segregation distortion. Since the genetic distance, when *GLO* is included, is very large compared with most examples of linkage disequilibrium, this might imply that the haplotype is maintained by suppression of recombination.

Extensive analysis is, however, necessary to confirm these points, and more convincing evidence that this haplotype represents a human *t*-haplotype could be obtained from studies of the segregation in a single large family. The linkage disequilibrium seen in the Boston population might be due to a non-specific co-selection ('hitch-hiker' effect) caused by the segregation distortion gene rather than suppression of recombination.

In a different type of study, linkage has been sought between genetic factors causing spina bifida and the MHC. A striking characteristic of several mouse *T/t* complex-associated mutations is their effects on the development of the axial skeleton, which in some cases are

similar to spina bifida in humans. In several studies, examination of large numbers of small families with affected individuals has not yielded an *HLA* association⁸⁻¹⁰ but this might be because there is more than one underlying cause. In the mouse, defects of this type are caused by several distinct unlinked genes, and may be occasionally due to environmental factors.

In apparent contrast, Amos and colleagues^{5,6} examined several generations of two families which appeared to be segregating for a dominant mutation that caused spinal defects and spina bifida occulta, although with varying degrees of severity in different individuals. The phenotype was found to show linkage with the MHC. Moreover, the *HLA* haplotypes, introduced by the original great grandparents, appeared to have an abnormal transmission ratio, suggestive of a human *t*-haplotype. A similar large family was found by Fellous *et al.*⁴ to have loose linkage between a hypothetical gene causing spinal defects and *HLA*. Again a human *t*-haplotype may be indicated.

Unfortunately, the conclusions of Amos

et al. and Fellous *et al.* remain equivocal because they depend very much upon the assumptions, such as dominance, penetrance and even segregation distortion, made in the model used for analysis. A much clearer result could be obtained by studying affected sib pairs, a technique which avoids applying assumptions about the mode of inheritance^{11,12}. Another problem is that the human mutants studied appear to be dominant. Although *H-2*-linked dominant mutations affecting the axial skeleton have been described in the mouse, the mutations associated with *t*-haplotypes isolated from wild mice are invariably recessive. Finally, segregation distortion may have several causes, for example, embryonic lethality. In the mouse, segregation distortion in males arises post-meiotically, but before embryogenesis, that is, the effect arises during the later stages of spermiogenesis, transport to the ovum or fertilization. The question has not been addressed in any of the human cases, although segregation distortion from male, but not female, heterozygotes is reminiscent of the mouse system. □

Developmental biology

Disorganized embryos?

from J.C. Smith

RECENT work has cast doubt on perhaps the most famous experiment in embryology. In 1924 Spemann and Mangold¹ showed, by transferring tissue between the gastrulae of differently pigmented species of newt, that transplantation of the dorsal lip of the blastopore to the ventral marginal zone caused the formation of a secondary embryo in which most of the somitic and lateral mesoderm, and almost the entire neural tube, were derived from the host. This dramatic result led Spemann to propose that the dorsal lip of the blastopore is the 'organizer' for the development of the whole embryo.

It has now become clear that the primary function of the organizer is to 'dorsalize' surrounding ventral mesoderm^{2,3} but most attention since 1924 has been concentrated on a more spectacular aspect of the organizer graft — the subsequent induction of the secondary nervous system. It did not take long to demonstrate that killed dorsal lips or even agar impregnated with tissue extracts or pure chemicals would induce patches of neuroepithelium from responding tissue but, despite this apparent early success, the biochemical basis of neural induction remains unknown to this day. It may be with this failure in mind that Marcus Jacobson⁴ has now suggested that 'neural induction' is nothing more than an artefact.

Since 1978 Jacobson has been studying the cell lineage of the central nervous system of the South African clawed toad, *Xenopus laevis*. The technique he uses is to

mark cells of the early *Xenopus* embryo by injecting a solution of horseradish peroxidase (HRP). The enzyme remains confined to the progeny of the injected cells and can be detected in frozen sections at later stages by standard histochemical techniques. By marking cells at progressively later stages of development and taking advantage of the cell mixing that occurs during gastrulation in *Xenopus*, Jacobson has concluded that the central nervous system (CNS) is composed of seven regions that he calls, by analogy with *Drosophila* development⁵, compartments.

Each compartment arises at the mid-blastula stage as a small group of founder cells. The descendants of each group of cells go on to contribute exclusively to one compartment and that compartment is not contributed to by the descendants of any other group of cells. Thus, if a cell is labelled after the establishment of the compartment founder groups its progeny will be restricted to one compartment; if it is marked earlier than this its progeny may contribute to more than one compartment.

What has this got to do with the organizer graft? Quite simply, if the pattern of the CNS is established at the mid-blastula stage, as Jacobson suggests, then it should not be possible to specify another CNS later on, during gastrulation. So

J.C. Smith is in the Laboratory of Experimental Morphology at the Mill Hill Laboratories of the Imperial Cancer Research Fund, Burtonhole Lane, London NW7 1AD.

1. Andrews, P.W. & Goodfellow, P.N. *Nature, News & Views* **299**, 296 (1982).
2. Artzt, K., Lockwood, M., Bennett, D., Kunz, H.W. & Gill, T.J. III *J. Immunogenet.* **9**, 371 (1982).
3. Awdah, Z.L., Raum, D., Yunis, E.J. & Alper, C.A. *Proc. natn. Acad. Sci. U.S.A.* **80**, 259 (1983).
4. Fellous, M. *et al. Am. J. med. Genet.* **12**, 465 (1982).
5. Amos, D.B., Ruderman, R.J., Mendell, N. & Johnson, A.H. *Trans. Proc.* **7** (Suppl. 1), 93 (1975).
6. Mendell, N.R. *et al. Cytogenet. Cell Genet.* **25**, 184 (1979).
7. Thomson, G. *Genetics* **85**, 753 (1977).
8. Bobrow, M., Bodmer, J.G., Bodmer, W.F. & McDevitt, H.O. *Tissue Antigens* **5**, 234 (1975).
9. de Bruyere, M., Kulakowski, S., Malchaire, J., Delire, M. & Sokal, G. *Tissue Antigens* **10**, 299 (1977).
10. Fellous, M., Hors, J., Boue, J., Dausset, J. & Jakob, F. *Birth Defects: Original Article Ser.* **15**, 3, 93 (1979).
11. Thomson, G. & Bodmer, W.F. in *HLA and Disease* (eds Dausset, J. & Sveigaard, A.) (Munksgaard, Copenhagen).
12. Spielman, R.S., Baker, L. & Zmijewski, C.M. *Ann. hum. Genet.* **44**, 135 (1980).

Jacobson went on to do some organizer grafts of his own, with the difference that the grafted organizers were taken from *Xenopus* embryos labelled with HRP. These grafts resulted in secondary embryos, as expected, but virtually all the cells of the secondary neural axis turned out to be labelled — that is, they were derived from the graft. Occasionally unlabelled cells were mixed with the grafted cells but in these cases, labelled cells were mixed with the host CNS as if exchange between graft and host cells had occurred. Thus, and in contrast with the results of Spemann and Mangold, Jacobson did not find induction of a secondary nervous system in *X. laevis*, but rather its development from the graft tissue itself.

What are we to make of this? Spemann and Mangold used, in their interspecies grafts, the urodeles *Triturus cristatus*, *T. taeniatum* and *T. alpestris*. *Xenopus laevis* is an anuran. Are we to believe that these two groups of amphibians differ in the origin of their nervous systems, as they do in the origin of their germ cells and in the location of the prospective mesoderm in the early gastrula^{6,7}? This would be unthinkable for those of us who believe in the universality of developmental mechanisms⁸ and who justify their experiments on amphibian embryos by arguing that the results will be applicable to all vertebrates. Perhaps an explanation may still be found in the exact details of the grafts performed and in the nature of the cell markers. Those who believe in the reality of neural induction will point out that a single positive experiment is enough to prove their case. Jacobson might well respond that interspecific grafts cannot provide the same degree of certainty as his own homoplastic transplants.

But while we wait for more results to settle this issue, it is worth repeating Slack's warning⁹ that to infer developmental mechanisms from fate map data alone, even fate maps as sophisticated as Jacobson's, is a dangerous business. The compartment hypothesis is now generally accepted in *Drosophila* not because of cell lineage studies but because of the existence of homeotic mutations, which convert one compartment into another. Until we see a frog with eyes in the back of its head the presence of compartments in vertebrates must remain open to question and criticism of Spemann and Mangold's experiment must remain uncertain. □

1. Spemann, H. & Mangold, H. *Arch. mikrosk. Anat. EntwMech.* **100**, 599 (1924).
2. Slack, J.M.W. & Forman, D. *J. Embryol. exp. Morph.* **56**, 283 (1980).
3. Cooke, J. *Am. Zool.* **22**, 91 (1982).
4. Jacobson, M. 'Origins of the Nervous System in Amphibians' in *Neuronal Development* (ed. Spitzer) (Plenum, New York, 1982).
5. Lawrence, P.A. & Morata, G. 'The compartment hypothesis' in *Insect Development* (ed. Lawrence) (Blackwell Scientific, Oxford, 1976).
6. Nieuwkoop, P.D. & Sutasurya, L.A. *J. Embryol. exp. Morph.* **35**, 159 (1976).
7. Smith, J.C. & Malacinski, G.M. *Dev. Biol.* (in the press).
8. Wolpert, L. *J. theor. Biol.* **25**, 1 (1969).
9. Slack, J. *Nature* **278**, 600 (1979).

Geophysics

A new way to estimate seismic source parameters

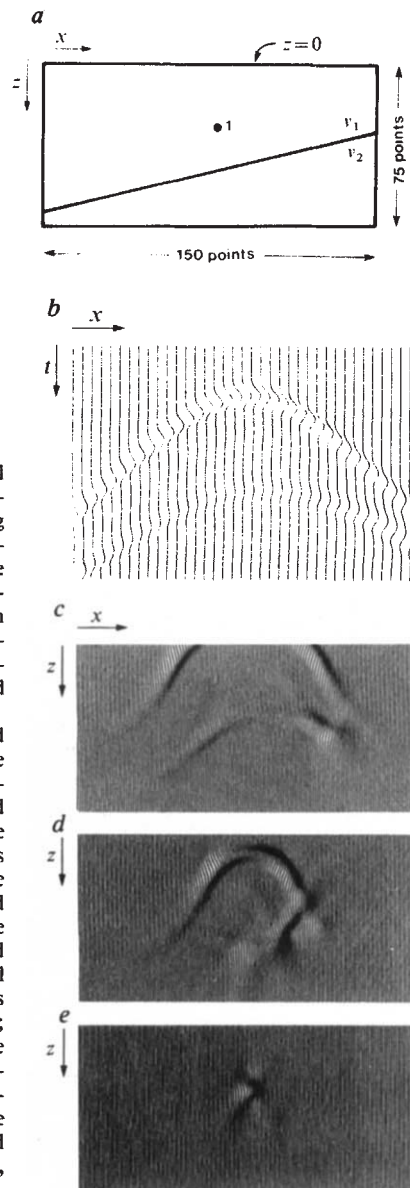
from Brian L.N. Kennett

SINCE the turn of the century, seismology has provided the bulk of the quantitative information about the Earth's interior. As instrumentation and theory have progressed, the aim has been to make the maximum use of the wavetrains recorded at each seismographic station. In studies of moderate-size earthquakes, using records 3,000–9,000 km away from the source, it is now common to calculate theoretical seismograms for each station, and to vary the assumed source parameters until good agreement is achieved between the theoretical and observed waveforms across the suite of stations. Large-scale use of this approach for close-in observations of small events has, however, been largely precluded by the quality of much seismographic data (rapidly recorded on smoked paper). Now the appearance of digitally recording small-scale seismic networks is making very different processing schemes possible.

A radical departure from the conventional approach has been suggested by McMechan (*Geophys. J. R. astr. Soc.* **71**, 613; 1982), who has adapted methods used in geophysical prospecting to the estimation of seismic source parameters. The reversibility of the wave equation is used to take the surface recordings at a dense network of receivers and extrapolate them back in time and space until a focus occurs at the source location.

In his demonstrations, McMechan used acoustic wave propagation (for both calculation of the seismograms and backward extrapolation) and a line of receivers. With a finite-difference treatment of the 'backwards' wave equation, the energy from the surface seismograms can be projected into a model of the subsurface. The process is illustrated in the figure.

The original model consists of a point source in a uniform region overlying a dipping reflector, which gives rise to a set of seismograms showing two distinct arrivals. The first pulse at each receiver is direct radiation from the source and the second pulse has been reflected from the



A single line source is excited in a two-dimensional model consisting of two uniform materials with differing acoustic wave speeds separated by a sloping boundary (a). The propagation of the waves is determined by a finite-difference solution of the wave equation and theoretical 'seismograms' are constructed at the surface $x=0$ as a sequence of records in time t at positions x . These records (b) show a symmetrical early arrival corresponding to direct propagation from the source and a later pulse reflected from the sloping boundary.

This set of surface 'seismograms' can then be used as the input to the source reconstruction scheme. The records are used as the boundary conditions for solving the wave equation running backwards in time and so concentrating energy near the original source point. Three stages in the reconstruction are shown as instantaneous 'snapshots' of the wave field in the original x - z plane of the model. In c, the reflected wave has entered the model and the later arms of the direct arrival have started to be propagated backwards. The central portion of the direct arrival has reached the subsurface in d and the wavefront is beginning to collapse into the source from above; meanwhile below the source position, waves are reflected back from the sloping interface. In e, the upper and lower wavefronts have coalesced into an image of the source. This is particularly sharp because the correct velocity model was used in the backward extrapolation. (From *Geophys. J. R. astr. Soc.* **71**, 613; 1982.)

interface. When these seismograms are used as the input to the spatial model incorporating the original velocity structure, the wavefronts concentrate with increasing extrapolation to give a sharp image of the source that generated them. It is also possible to distinguish multiple sources occurring at different times and so, in principle, detect the propagation speed of a source.

Improved estimates of seismic source locations would be particularly useful in several problems. Following a major earthquake there is normally a sequence of smaller 'aftershocks', and the spatial distribution of these events often provides the best indicator for the region of the fault ruptured in the original earthquake. Also, the changing location of seismic events beneath a volcano can give information on the movement of magma and the likelihood of an eruption.

How far, then, is a scheme such as McMechan's able to help with improving source parameters? With a scattered network of stations it is clear that no improve-

ment is likely over existing methods. A tightly distributed network of digital stations would give good results, provided the seismic properties of the region are well established. Imperfections in structural information would be reflected in smearing of the source images and the generation of 'ghost' sources. A potential advantage of the approach is that no specific interpretation needs to be made of the various arrivals on the seismograms, since their interrelation is unravelled during the backward extrapolation. But, when both compressional waves and transverse waves are recorded at the same station for a particular event, it will require recordings of all three components of displacement and the use of the full elastic wave equation to get the best image of the source.

The new method is likely to stimulate much further work but offers no panacea for painless seismic source extraction. □

Brian L.N. Kennett is in the Department of Applied Mathematics and Theoretical Physics, Silver Street, Cambridge CB3 8EW.

and longer periods.

Analysis of seismic waves by Choy and Cormier allowed them to conclude that there are strong gradients in the velocities of P and S waves and anelasticity in the upper 200–300 km of the inner core. The compressional velocity jump is about 5 per cent, substantially less than previous estimates, and the shear velocity may be nearly zero at the top of the inner core. Another recent study⁵ presents models with high gradients in shear velocity over the outer half of the inner core. These results plus the high gradient in bulk modulus and anelasticity are suggestive of a broad transition region.

The pressure interval of the anomalous gradient is 0.1–0.3 megabars, much greater than the pressure interval of transition regions in the mantle. It is hard to imagine how a fluffy or partially molten or two-phase inner core could exist over such a large pressure interval. Turbulent convection could keep solid particles in suspension, but this would require a high temperature gradient. The temperature gradient at the center of the Earth, of course, should be zero. It is also hard to reconcile the high attenuation of the outer part of the inner core with a slurry. In this case the Q should initially decrease with depth, starting at the high values appropriate for the fluid outer core.

If the inner core is a viscous fluid, with the relaxation time increasing with pressure, then the shear velocity, for a given frequency seismic wave, increases from zero to a finite value controlled by the viscosity. The absorption is very high where the velocity is low and increases as the velocity increases. If this is true, the transition depth would depend on frequency and the boundary would have little to do with the melting point of iron. A viscous fluid can transmit shear waves but would still be able to flow in response to longer-term stresses. The interesting possibility would then exist that there is no inner core at all on time scales relevant to convection in the core.

The concepts of a solid and crystallizing inner core are important in calculations of the geometry and energy sources of the dynamo. If the inner core is a viscous fluid, rather than a crystalline solid, then the anomalously low shear velocity and high Poisson's ratio would have a simple explanation. A critical experiment now is to determine whether the radius of the inner core varies with frequency. □

Earth science

A new look at the inner core of the Earth

from Don L. Anderson

THE inner core is slightly smaller than the Moon but is more than three times denser. It represents less than two per cent of the mass of the Earth but nevertheless has had important effects on the Earth's magnetic field and on the thermal history of the core.

The change in physical properties at a radius of about 1,215 km that defines the inner core has usually been interpreted as stemming from a melting phenomenon related to crystallization of the inner core. A recent high-resolution study by Choy and Cormier¹ of the outer part of the inner core indicates that this may actually be an oversimplification. The rigidity and viscosity of the outer core are known to be very low and convection is presumed to take place with ease, as this is required by dynamo theories of the Earth's magnetic field. The new study suggests that the rigidity at the top of the inner core is also very low, perhaps zero, and increases gradually with depth.

The inner core was discovered in 1936 by the Danish seismologist Inge Lehmann². Birch³ and Bullen⁴ concluded that it was solid, and later work showed that its density was consistent with that of solid iron with, perhaps, some nickel. Thus, the inner core appeared to be a small dense solid body suspended in a fluid outer core.

More recently it has been shown that at

least part of the inner core–outer core boundary is sharp, that its Poisson's ratio is very high and that it attenuates seismic energy more rapidly than the outer core. Oscillation and precession of the inner core may be important in maintaining and modifying the dynamo. It is not clear whether the inner core shares the same rotation axis as the mantle, whether it spins at the same rate, or if it follows the polar wanderings of the mantle.

It is usually assumed that the inner core–outer core boundary represents the zone where the melting point of the iron-rich alloy making up the outer core is passed. Although the inner core may simply represent frozen outer-core material, no light element is required in its makeup. If the inner core has crystallized over time, it may have excluded the lighter elements to the outer core. The inner core may also represent a glassy-type transition^{6,7} and simply have a higher shear viscosity than the outer core rather than have the rigidity of lattice. If this is the case, the increase of 'rigidity' with depth would be gradual, since pressure can be expected to increase viscosity. The 'boundary' between the outer and inner cores would be the depth at which the viscous relaxation time is the same as the frequency of the seismic wave which is interrogating it. The location of the boundary would be frequency dependent. The inner core would be much smaller at free oscillation periods than for 1-s body waves, and perhaps would not exist at tidal

Don L. Anderson is in the Seismological Laboratory, California Institute of Technology, Pasadena, California 91125.

1. Choy, G.L. & Cormier, V.F. *Geophys. J. R. astr. Soc.* **72**, 1 (1983).
2. Lehmann, I. *Publ. Int. Geod. Geophys. Un., Ass. Seismol. Ser. A., Trav. Sci.* **14**, 87 (1936).
3. Birch, F. *Am. J. Sci.* **238**, 192 (1940).
4. Bullen, K.E. *Nature* **157**, 405 (1946).
5. Häge, H. *Phys. Earth planet. Inter.* **31**, 171 (1983).
6. Gutenberg, B. *Bull. seism. Soc. Am.* **48**, 301 (1958).
7. Anderson, D.L. *Nature* **285**, 204 (1980).

Contribution no. 3897, Division of Geological and Planetary Sciences, California Institute of Technology.

Mathematics

High-speed tests for primes

from Ian Stewart

PRIME numbers have exercised the imagination and ingenuity of mathematicians since at least the time of Euclid. The recent upsurge of interest in public-key encryption¹ has added a practical, indeed military, dimension to two venerable questions: first, given a number, is it prime or composite, and second, if it is composite, what are its factors?

In principle, there exist specific procedures, or algorithms, that can solve both problems: if the number is n , try all possible divisors up to $\sqrt{n}$. In practice, these are hopelessly inefficient, taking perhaps a million years to test a 40-digit number on a fast computer. Better though less direct methods exist, and until recently the best techniques might solve the first question for an 80-digit number and the second for a 40-digit number in an acceptable time. This reflects a general (but conjectural) view that the second problem is substantially harder than the first.

The efficiency of an algorithm is usually gauged by estimating how rapidly the computational time grows as the number of digits k of the number n increases. If this time is always smaller than Ck^r for some constant C and a fixed power k^r of k , then the algorithm is said to run in polynomial time. Trial-and-error division runs in exponential time $C2^k$, which (as Thomas Malthus vividly insisted when $r=1$) grows infinitely faster in the long run.

For a truly practical algorithm the value of the constant C is also important. For instance, an algorithm that required exactly one billion years, for any n , would run in polynomial (indeed constant) time. But in less artificial cases, a polynomial-time algorithm is usually efficient and practicable, whereas an exponential-time algorithm is not.

Whether the problem of deciding whether a number is prime or composite can be solved in polynomial time remains in doubt. (It is widely believed that the factors of a composite number cannot be found in polynomial time, but this has never been proved.) The current status of this problem is a catalogue of useful near-misses.

The basis of most modern primality tests is Fermat's 'Little' Theorem: if p is prime and a is not divisible by p , then $a^{p-1} - 1$ is divisible by p . This can be used to prove a number composite, without exhibiting any factors. For example, when $p=4$ and $a=2$, one must evaluate $2^4 - 1 = 15$. This is not divisible by 4, so 4 cannot be prime. By applying this test for a suitable number of randomly chosen numbers a , it is possible to obtain a probabilistic algorithm for

primality that runs in polynomial time. 'Probabilistic' means that the procedure terminates within the time limit with a probability near 1, and if it does terminate, will prove or disprove primality.

Unfortunately, some composite numbers satisfy the Fermat Theorem for all a , 'by accident'. They are called Carmichael numbers, and an example is the number $561 = 3 \cdot 11 \cdot 17$. Refining the Fermat Test eliminates some, but not all, of these: the numbers remaining are called strong pseudoprimes to the base a .

A composite number cannot be a strong pseudoprime to every base; so the proper selection of bases might lead to an efficient algorithm for primality. In 1976, G. Miller² showed that the choice of bases can be made so as to yield an algorithm that runs in polynomial time — but only by assuming the validity of the 'Extended Riemann Hypothesis', one of the most famous and difficult unsolved conjectures in mathematics.

In 1980 L. Adleman and R. Rumely announced a 'nearly' polynomial-time algorithm, which has recently been published³. It involves more general strong pseudoprimal tests, using bases a that are of the form

$$a_0 + a_1 \xi + \dots + a_{m-1} \xi^{m-1}$$

where ξ is a complex m th root of unity. These cyclotomic integers were introduced by Ernst Kummer in about 1843 in an attack on another famous conjecture, Fermat's Last Theorem (the equation $x^p + y^p = z^p$ has no integer solutions for $p > 3$) and led to the modern theory of algebraic numbers. By exploiting some deep theorems from algebraic number theory, such as the 'Higher Reciprocity Laws', Adleman and Rumely were led to conjecture a running-time for their algorithm of the form

$$k^C \log \log k$$

intermediate between polynomial and exponential time. Using additional number-theoretic results, Carl Pomerance and Andrew Odlyzko proved this conjecture. The algorithm, at least in an improved form, is computer-practical: estimates suggest that 200-digit numbers may eventually be tested for primality in a few minutes on a fast computer.

Number Theory is usually thought (by outsiders) to be one of the purest and least applicable areas of mathematics. The work of Adleman, Rumely, Pomerance, Odlyzko and others suggests that a major revision of this view may be in order. And it is likely that many other questions about the efficiency of algorithms will also require number-theoretic methods of analysis. □

1. Rivest, R., Shamir, A. & Adleman, L. *Commun. ACM* **21**, 120 (1978).
2. Miller, G.L. *J. Comput. Systems Sci.* **13**, 300 (1976).
3. Adelman, L.M., Pomerance, C. & Rumely, R.S. *Ann. Math.* **117**, 173 (1983).

Astronomy

Quasars: searching out the smallest and nearest

from Martin Elvis

MUCH of the immediate excitement in astronomy comes from finding the biggest, the most luminous or the most distant of the various kinds of object. Particular satisfaction for those who search out extremes is provided by quasars, for they can be seen out to redshifts of 3.5 — equivalent to looking back about 80 per cent of the age of the Universe. When it comes to studying the details of quasars' physical structure, and thus to understanding how they work, there is a problem, however. The smallest dimension that can be directly resolved using current technology at a redshift of 3 is about 15 pc (10^{19} cm). This is remarkably small but, unfortunately, it is still not small enough. Variability in quasars and other active galaxies is seen on time scales of a few days¹, implying that the radiating region must be at least a few light-days across — about 10^{16} cm.

To study quasar structure in detail, we have to change our perspective and look

for much nearer, but weaker, examples. In the past two years, several of the nearest galaxies have been found to contain miniature examples of quasars in their nuclei. Now radio observations² have begun to examine these miniature quasars at just this scale. In the first miniature quasar examined, an elongated structure aligned with the minor axis of the galaxy was found. This suggests that the tiny central quasar 'knows' the orientation of the much larger galaxy.

Norbert Bartel of MIT and his collaborators at Haystack Observatory, JPL and the Max-Planck-Institut in Bonn used very long-baseline interferometry (VLBI) to achieve this result. It is the first time that structure has been seen in an active nucleus (the term used for low-luminosity quasars) close to the scale of the intense radiation

Martin Elvis is in the Smithsonian Astrophysical Observatory, Cambridge, Massachusetts 02138.

Ian Stewart is in the Mathematics Institute of the University of Warwick, Coventry CV4 7AL.

emission. Such observations have become possible only through the introduction of the new Mark III VLBI system, which is five times more sensitive than its predecessor³.

The galaxy studied in the first Mark III observations was the large spiral Messier 81 (M81). It has long been known that M81 has a compact nuclear radio source in its nucleus, on the basis of both variability⁴ and a simple detection as a VLBI source⁵. More recently, it was found to have an optical spectrum similar to that of quasars⁶ and a central X-ray source characteristic of quasars⁷. Since quasars and radio galaxies often show very linear jet-like radio structure⁸, it was natural to choose M81 for study as soon as the Mark III VLBI system became available.

The first results of Bartel *et al.* show a clearly elongated structure in M81, about 6×10^{16} cm long and strikingly co-aligned with the apparent minor axis of the spiral galaxy as seen in the sky. This is what would be expected if the radio emission were showing a jet coming out of the galaxy perpendicular to the galaxy disk, although this is not a necessary deduction. The alignment would suggest that the tiny active nucleus (with a size only 10^{-6} that of the

galaxy) 'knows' the rotation axis of the much larger object in which it is embedded. Some physical mechanism is needed to link these two different scales. One direct connection might be through the accreting material that is commonly thought to power the activity. If the material originates in the galaxy disk, then it will conserve its angular momentum vector as it spirals into the nucleus. With a sample of only one galaxy so far, however, it is not yet possible to know whether the alignment is real or is a random coincidence. On a larger scale (10^{21} cm) VLA data do not show any alignment for more powerful active nuclei in spirals⁹.

The first Mark III data are only a beginning. Already the same group of astronomers has data in hand using many

more telescopes to give more sensitivity and more baselines which will allow a real map of the M81 source to be constructed. (The current work had to rely on model fitting.) The extra sensitivity should allow an extension of the work to other, not quite so nearby, galaxies. Then we shall be able to discover whether the alignment is accidental or indicates a real link between the active nucleus and the parent galaxy.

Perhaps the most fascinating prospect in the near future is to repeat the observations at higher frequencies. By working at wavelengths as short as 1.3 cm the VLBI technique can map details three times as fine, which would correspond to 8×10^{15} cm in M81. This is not only the scale on which accretion disks should be found — if they exist in quasars — it also opens up the possibility of seeing motions in the nucleus on a time scale of a few years. Both rotation and expansion can be searched for. Optical work shows that mass motions of gas at speeds of $5,000 \text{ km s}^{-1}$ exist in M81 (ref. 6). If these motions relate to the radio source and were mainly in expansion, as some models predict¹⁰, then changes in the source size should be detectable with high-frequency VLBI over a time span of 3–5 years. □

1. Marshall, N., Warwick, R.S. & Pounds, K.A. *Mon. Not. R. astr. Soc.* **194**, 987 (1981).
2. Bartel, N. *et al. Astrophys. J.* **162**, 556 (1982).
3. Rogers, A.E.E. *et al. Science* **219**, 51 (1983).
4. Crane, P.C., Gioffrida, T.S. & Carlson, J.B. *Astrophys. J. Lett.* **203**, L113 (1976).
5. Kellerman, K.I. *et al. Astrophys. J. Lett.* **210**, L121 (1976).
6. Peimbert, M. & Tores-Peimbert, S. *Astrophys. J.* **245**, 845 (1981).
7. Elvis, M. & Van Speybroeck, L. *Astrophys. J. Lett.* **257**, L51 (1982).
8. Readhead, A.C.S. & Pearson, T.J. *IAU Symp.* **97**, 279 (1982).
9. Wilson, A. *IAU Symp.* **97**, 179 (1982).
10. Mathews, W. *Astrophys. J.* **252**, 39 (1982).



100 years ago THE ORNITHOLOGIST IN SIBERIA

The ornithologists are certainly among the most enterprising of the seekers after truth. John Gould, the Birdman, is dead, but the same spirit which led him over the seas fifty years ago to investigate the then unknown Ornis of Australia still animates his brother bird-men. Mr. Henry Seebohm — a distinguished Member of the British Ornithologists' Union — has recently made two journeys into Northern Siberia, solely with the object of observing new forms and habits of bird-life and of collecting specimens. The scientific results of these expeditions have been published in the *Ibis* — the organ of the British Ornithologists' Union —

which is now entering upon the twenty-fifth year of its existence.

The first of these two expeditions, to the lower valley of the Petchora, in North-Eastern Russia, was made by the author in 1875. For the second Mr. Seebohm left London on March 1, and travelled by rail to Nishni Novgorod, a distance of some 2,400 miles. Thence was a sledge-journey of about 3,200 miles to the winter quarters of the good ship *Thames*, on the Yenesay, or rather a little way up the Koorayika, an affluent of the Yenesay, on its right bank. The crew of the *Thames* who had passed a long and dreary winter, frozen up at this spot, were found on the travellers' arrival to be well and hearty, owing to the judicious precautions that had been taken by their Captain for the benefit of their health.



FIG. 3.—Driving with the ice on the Koorayika.



FIG. 2.—Little Stint's nest, eggs, and young.

A comparison of volcanic eruption processes on Earth, Moon, Mars, Io and Venus

Lionel Wilson*† & James W. Head III*

* Department of Geological Sciences, Brown University, Providence, Rhode Island 02912, USA

† Lunar and Planetary Unit, Department of Environmental Sciences, University of Lancaster, Lancaster LA1 4YQ, UK

The silicate planets and satellites display a wide range of physical, chemical and atmospheric characteristics which may influence the nature of volcanism, a major geological process common to the evolution of the surfaces of these bodies. Consideration of the process of magma ascent and eruption from first principles allows predictions to be made concerning volcanic eruption styles and expected landforms and deposits on each planetary body. Examination of actual landforms and deposits in light of these predictions leads to a better understanding of the nature of volcanic eruption processes and outlines outstanding problems.

THE investigation and comparison of the surfaces of the Earth and the other silicate planets during the past two decades have revealed the importance of volcanism, especially basaltic volcanism, in the formation and evolution of the crusts of these bodies¹⁻⁵. On Venus, the nature of volcanic deposits is unknown but volcanism is thought to have an important role in the evolution of the planet^{6,7}. While there are many similarities between the styles of volcanic activity which are presently observed or have occurred in the past on the Earth, Moon, Mars and Io, there are also significant differences. Active eruptions currently occur on Earth and Io, while little or no volcanic activity seems to have occurred on the Moon in the latter half of Solar System history. On the Moon we see sinuous channels 10 times larger than terrestrial lava channels and on Mars, several shield volcanoes many times larger than Hawaii are observed.

How do we account for these and other differences? Eruption style is controlled by a combination of factors such as magma composition and temperature (which control magma rheology), magma volatile content (which tends to be the major influence on eruption velocity in the vent), tectonic setting (which influences the depth from which magma erupts) and external environment (such as atmospheric pressure and gravity). The Earth presents a rich array of volcanic landforms representing magmas of different compositions, eruptions through oceanic and continental crusts, and varying tectonic environments associated with different types of lithospheric plate boundaries. It may, therefore, be of fundamental importance that the smaller terrestrial planetary bodies (Moon, Mars, Mercury) are characterized by a single unsegmented and stable global lithospheric plate^{1,2}.

Specific studies of the ascent of magma to the surface of a planet have, until recently, concentrated on the Earth and Moon and have covered: the modelling of the diapiric rise of large magma bodies through the lower crust⁸⁻¹⁴; mechanical¹⁵⁻²⁰ and thermal^{16,20-22} aspects of the emplacements of dykes and sills; static²³ and dynamic^{24,25} aspects of the flooding of the lunar mare basins; and the dependence of the rise speeds and discharge rates of steadily erupting magmas²⁵⁻³⁰ on their rheological properties and volatile contents and on their thermodynamic properties^{29,31,32}. Recent work has extended the latter studies to Mars^{33,34} and Venus^{35,36} and has allowed treatments of the basic mechanisms of explosive eruptions^{37,38} and of the interaction of eruption clouds with the atmosphere³⁹⁻⁴¹ to be generalized to other planetary environments. The discovery⁴² of active volcanism on Io, innermost of Jupiter's galilean satellites, has

emphasized the need to develop treatments of eruptions on bodies with negligible atmospheres⁴³⁻⁴⁷. It has also prompted the need to study the processes occurring in bodies where tidal dissipation⁴⁸, rather than radiogenic heat production, dominates the thermal budget. We now review the basic principles of magma ascent and eruption and discuss the relationship of these principles to eruption style. We conclude with a discussion of eruption styles and the predicted and observed volcanic landforms and deposits on the Earth, Moon, Mars, Venus and Io.

Magma ascent through the crust and eruption

The rise velocity, u_d , of an erupting magma at depths below the surface sufficiently great that volatile exsolution is negligible can be specified in terms of the magma density, ρ , viscosity, η , and yield strength, y , the fissure width, w , the planetary gravity, g , and the effective density difference, $\Delta\rho$, between the rising magma and the surrounding country rock. This effective density difference will be equal to (or slightly less than) the true density difference in the case of a magma utilizing a fissure opened to the surface by external tectonic forces. In the case of a magma erupted from a reservoir in which an excess (non-hydrostatic) pressure has accumulated and fractured the overlying rocks⁴⁹, $\Delta\rho$ can be several times greater than the true difference. u_d is given by

$$u_d = (w^2 g \Delta\rho - 2yw) / 12\eta \quad (1)$$

over a very wide range of conditions²⁵ and the corresponding mass eruption rate, $\dot{M}$, is given by

$$\dot{M} = wL\rho u_d = w^2 L\rho (wg\Delta\rho - 2y) / 12\eta \quad (2)$$

where L is the fissure length (measured in the horizontal plane). The strong dependence of $\dot{M}$ on the fissure width, w , means that the effect on effusion rate of differences in tectonic environments (which control the upper limits on w and L through strain accumulation) or depths of origin of magma (which control the lower limit on w through magma cooling constraints^{18,25}) between planets will be much more important than the effects of differences in gravity and may be at least comparable with the effect of differences between the rheological properties of various magma types on a single planet. Rheological parameters vary by factors of $\sim 10^4$ (ref. 25) and at least 10^3 and 10^2 among terrestrial, lunar and martian magmas, respectively. The lunar and martian values are deduced from observations of the morphologies of lava flows^{50,51} visible in



Fig. 1 Lava flows on the flanks of Arsia Mons, Mars. Individual flows average 5 km width and have been traced over distances of 300 km but may extend many hundreds of kilometres further. Viking Orbiter frame 56A26.

spacecraft pictures (Fig. 1) and so may be biased towards the more readily visible products of large-volume eruptions of low-viscosity magma; geochemical arguments suggest⁵² that silicic magmas may have been produced on Mars, but if their products have taken the form of localized, viscous lava flows⁵³ or domes⁵⁴, they may not be readily visible. Nevertheless, the great lengths of many sinuous rilles^{55,56} and lava flows⁵⁷ on the Moon may be interpreted⁵⁷⁻⁵⁹ to imply that lunar basaltic eruption rates commonly lay in the range 10^6 – 10^9 kg s⁻¹ near the close of the era of mare basin filling ($\sim 3,100$ Myr ago). These values may be compared with eruption rates in the range 10^2 – 10^7 kg s⁻¹ (refs 25, 60) for quaternary basaltic eruptions on Earth. The higher lunar values imply²⁵ fissure widths a factor of ~ 3 times larger than those commonly found on Earth. Many lunar source vents are associated with major impact basin ring systems and the generally extensional state of stress in the lithosphere in early history may have favoured maintaining open fractures. In later lunar history, the high effusion rates may well be a reflection of the relatively great thickness of the lunar elastic lithosphere implied by thermal models²³. Stratigraphically young lava flows on Mars also exhibit lengths in excess of 100 km (refs 4, 61–63) and may similarly be related to the presence of a thick lithosphere.

As an erupting magma approaches a planetary surface, the external pressure decreases and volatiles begin to exsolve into nucleating bubbles. The expansion of the resulting gas (at a fairly constant temperature, buffered by the heat content of the liquid magma) releases energy which causes the ascent velocity to increase. If the growing gas bubble volume fraction in the melt exceeds about 3/4, the magma may disrupt to form pyroclastic fragments⁶⁴ and an explosive eruption will result; if disruption does not occur, the eruption will be effusive, producing a vesicular lava flow.

Effusive eruptions

As erupted magma emerges from a vent or accumulates at the base of a fire fountain it forms a lava flow moving downslope from the source under gravity. The initial morphology of the flow is determined by the rheology of the lava^{50,65}, which is usually modelled as a Bingham plastic with properties controlled by two parameters, the yield strength and the plastic viscosity. The finite yield strength dictates that lava motion can only occur in a central channel contained between stationary

levees. If the levee width is w_b on each side of a flow moving on a slope α and the maximum thickness of the levee is d_c , it can be shown that⁵⁰

$$d_c = y/g\rho\alpha \quad (3)$$

$$w_b = y/2g\rho\alpha^2 \quad (4)$$

Thus, lava flows will tend to be thicker and to have wider levees on low-gravity planets, other factors being equal. If the width of the central channel of a flow is w_c and the variable $R = w_c/2w_b$ is introduced, it is found that the volume of lava erupted per second, F , can be related to the flow geometry by (ref. 56 and L.W., unpublished calculations)

$$F\eta(g\rho)^3(\alpha/y)^4 = \begin{cases} R^3/24, & 0 < R \leq 1 \\ R^{11/4}/24, & R \geq 1 \end{cases} \quad (5)$$

which invert to

$$w_c = \begin{cases} (24F\eta/y\alpha^2)^{1/3}, & 0 < R \leq 1 \\ (24F\eta)^{4/11} g^{1/11} \rho^{1/11} y^{5/11} \alpha^{6/11}, & R \geq 1 \end{cases} \quad (6)$$

showing that central channel width is essentially independent of gravity⁵⁶. The average flow velocity in the central channel, u_a , is found by inserting equations (3) and (6) into the continuity equation, $F = u_a w_c d_c$:

$$u_a = \begin{cases} F^{2/3} g \rho \alpha^{5/3} / 24^{1/3} \eta^{1/3} y^{2/3}, & 0 < R \leq 1 \\ F^{7/11} g^{10/11} \rho^{10/11} \alpha^{17/11} / 24^{4/11} \eta^{4/11} y^{6/11}, & R \geq 1 \end{cases} \quad (7)$$

and so is essentially proportional to the gravity, other factors being equal.

Moving flows lose heat by conduction to the substrate, by radiation to the surroundings and by convection and conduction to the atmosphere if one is present. A thermal boundary layer, within which η and y increase rapidly, grows inwards from all surfaces of the flow⁵⁶, and ultimately motion ceases. In some cases the upper surface of a flow becomes rigid long before the interior and motion takes place in an enclosed (and hence thermally insulated) tube within the body of the flow. There is theoretical⁵⁷ and observational^{58,66} evidence for the suggestion that flow ceases in terrestrial lava channels when the (dimensionless) Gratz number, $Gz = Fd_c/\kappa x w_c$, has decreased from an initially large value to about 300; here κ is the thermal diffusivity and x the distance flowed from the vent. The maximum length, X , of a single flow should therefore be (L.W., unpublished calculations)

$$X \approx Fd_c/300 \kappa w_c = \begin{cases} F^{2/3} y^{4/3} / 865 \kappa \eta^{1/3} \alpha^{1/3} g \rho, & 0 < R \leq 1 \\ F^{7/11} y^{16/11} / 953 \kappa \eta^{4/11} \alpha^{5/11} g^{12/11} \rho^{12/11}, & R \geq 1 \end{cases} \quad (8)$$

The left-hand expression above shows why there is some terrestrial field evidence^{58,67} for flow length being proportional to effusion rate while other data⁶⁸, presumably involving flows with wider ranges of values of d_c and w_c , do not show this simple trend so clearly. The critical value of Gz for flow motion to cease may not be the same for all planetary environments; for example, the high ambient temperatures (~ 700 K) on Venus would lead to significantly slower cooling rates and, hence, greater lengths for lava flows on that planet³⁶ (by a factor of the order of 2–3 in both cases). The right-hand expressions above indicate that, all other factors being equal, flows should be longer on low gravity planets in inverse proportion to the gravity. The dependence of X on gravity is less striking, however, if it is recalled that F also involves the value of g : if $F = M/\rho$ is substituted from (2) (and $wg\Delta\rho$ is assumed to be much greater than $2y$, which is generally true) it is found that X is proportional to $w^2/g^{1/3}$ ($0 < R \leq 1$) or $w^{21/11}/g^{5/11}$ ($R \geq 1$) which emphasizes again the strong dependence of flow length on crustal fissure width.

All of the above equations essentially assumed that lava moves in a laminar fashion; however, turbulent motion can occur⁵⁵ if F exceeds a critical value F_t given by (L.W., unpublished calculations)

$$F_t = \begin{cases} 1,113y^{1.3}\eta^{0.8}/\rho^{2.1}\alpha^{2.2}g^{1.2}, & 0 < R \leq 1 \\ 2,053y\eta/\rho^2\alpha^2g, & R \geq 1 \end{cases} \quad (9)$$

On a slope of 0.01 rad, this critical value of F corresponds to a mass eruption rate of $M_t \approx 3 \times 10^7 \text{ kg s}^{-1}$ for both lunar⁵⁷ and terrestrial (unpublished data) basalts. It is rare⁶⁰ for terrestrial basaltic eruption rates to exceed $3 \times 10^6 \text{ kg s}^{-1}$, whereas lunar basalts seem to have commonly been extruded at rates between 10^8 and 10^9 kg s^{-1} (refs 57, 59), and the turbulent nature of some of these lavas has led to the formation of sinuous rille channels by thermal erosion^{55,56,59}. Similar channels probably exist on Mars³, but no such features are seen on Mercury; the resolution of radar images of Venus is, as yet, inadequate to resolve such features even if they are present.

The relationships developed above are only applicable to liquids having rheological parameters similar to those of a Bingham plastic, and their application to lava flows is generally complicated⁶⁵ by such factors as variations in effusion rate from the vent (causing overflow or enlargement of the initial levees) and accretion onto the levees of cooled clots of lava at the flow surface. The equations will not apply at all to the flows of liquid sulphur observed⁶⁹ on Io from the Voyager spacecraft. The apparent viscosity of sulphur displays a local minimum at about 430 K (ref. 69) and so a completely new treatment of the motion and cooling characteristics of sulphur flows is required^{70,71}.

Steady explosive eruptions

When the initial volatile content of a steadily erupting magma is high enough to cause disruption of the magma into pyroclasts⁶⁴ (scoriaceous in mafic magmas and pumiceous in silicic melts) and released gas, the eruption products emerge from the vent at high speeds since the effective viscosity of the gas-clast mixture⁷² is several orders of magnitude less³⁰ than that of the liquid magma. If a magma contains a total volatile weight fraction n_t before exsolution begins and the solubility of the volatile species as a function of the pressure, P , is $n_s(P)$ (refs 25, 64), then the exsolved volatile weight fraction is

$$n_e = n_t - n_s(P) \quad (10)$$

The volumes of gas and liquid are proportional to $(n_e QT/mP)$ and $(1-n_e)/\rho_1$, respectively³⁰ where ρ_1 is the magma liquid density, Q is the universal gas constant and T and m are the temperature and molecular weight of the gas, respectively. If

the magma/bubble foam disrupts when the bubbles occupy 3/4 of the volume⁶⁴, we have

$$3(1-n_e)/\rho_1 = n_e QT/mP \quad (11)$$

and if n_e is eliminated from equation (11) with equation (10), a value can be found for the pressure at which disruption occurs, P_d . If the value found for P_d is less than the atmospheric pressure of the planet, P_a , this implies that bubble growth is insufficient to disrupt the magma. When magma disruption does occur, P_d can be inserted into equation (10) to give the value of n_e at disruption, say n_d . Numerical simulations then show^{25,30,38} that the final velocity of the erupted gas/clast mixture, u_t , is well approximated by

$$\frac{1}{2}u_t^2 = \frac{1}{2}u_d^2 + K \frac{n_d QT}{m} \log_e (P_d/P_a) \quad (12)$$

where K is a factor representing frictional and potential energy losses and is commonly only slightly less than unity³⁸. Table 1 gives some values of u_t as a function of n_d for eruptions on the Earth, Mars and Venus for both H_2O and CO_2 as the volatile phase. Eruption speeds are seen to be about 1.5–2 times larger on Mars than on Earth due to the lower value of P_a ; but the drag force on ejected pyroclasts, which determines the size of the largest particle which can be ejected and is proportional to the product of erupted gas density and u_t^2 , is less than in the terrestrial case as the gas density, being proportional to P_a , is roughly 150 times smaller. The largest pyroclasts commonly ejected in martian explosive eruptions should have been only several centimetres in size, in contrast to sizes up to a few metres observed on Earth⁷³. Similar arguments for Venus indicate that the largest pyroclasts produced there could be up to 20 m in size⁷⁴.

Magma disruption and consequent explosive activity can only occur (Table 1) if magma volatile contents exceed ~0.01 wt% on Mars, ~0.07 wt% on Earth and ~3 wt% on Venus, the exact values being dependent on volatile species and solubility and on atmospheric pressure. The presence of abundant CO_2 and H_2O (as ice) on Mars^{75,76} argues for extensive planetary degassing via eruptions which were commonly explosive, irrespective of magma composition. Many of the larger lava flows, which morphology suggests have a basaltic composition^{4,51}, were probably fed by coalescing clots of uncooled magma falling from optically thick fire fountains²⁵ as were the flows that formed the lunar sinuous rilles^{33,59}. The need for volatile contents in excess of ~3 wt% (uncommonly high for Earth) to permit significant explosive activity on Venus may imply that such activity is not, or has not been, common there³⁶. However,

Table 1 Minimum mass eruption rate, M_c , required to ensure eruption column collapse to form pyroclastic flows and eruption velocity, u_t , as a function of exsolved H_2O and CO_2 content, n_d

H_2O n_d (wt%)	Earth		Mars		Venus lowland		Venus highland	
	u_t (m s^{-1})	M_c (kg s^{-1})	u_t (m s^{-1})	M_c (kg s^{-1})	u_t (m s^{-1})	M_c (kg s^{-1})	u_t (m s^{-1})	M_c (kg s^{-1})
5	530	2.2×10^9	770	1.2×10^9	267	2.7×10^9	205	1.4×10^9
3	380	5.3×10^8	580	1.7×10^8	170	3.0×10^8	90	3.0×10^7
1	200	4.8×10^7	330	1.5×10^7	—	—	—	—
0.3	90	2.8×10^6	180	2.1×10^6	—	—	—	—
0.1	40	1.4×10^5	80	1.0×10^5	—	—	—	—
0.03	—	—	25	1.0×10^3	—	—	—	—
CO_2								
5	305	2.2×10^8	464	4.8×10^7	105	4.8×10^7	—	—
3	240	6.0×10^7	340	1.4×10^7	—	—	—	—
1	109	5.1×10^6	190	1.8×10^6	—	—	—	—
0.3	45	2.0×10^5	85	1.0×10^5	—	—	—	—
0.1	15	2.5×10^3	40	7.0×10^3	—	—	—	—
0.03	—	—	20	7.0×10^2	—	—	—	—

Values are given for eruptions at the 3 mbar level on Mars (appropriate to the upper parts of the Tharsis volcanoes) and at the 50 and 90 bar levels on Venus (appropriate to highland and lowland eruption sites, respectively). A blank entry means that magma disruption to form pyroclasts cannot occur.

material at the Venera 13 landing site, on the flanks of Beta Regio, has a major element composition⁷⁷ very similar to that of terrestrial alkali basalts commonly found to be rich in CO₂ and halogens. This may mean that some explosive activity can occur, at least locally.

On any planet with an appreciable atmosphere (such as Earth, Mars or Venus), an explosively ejected cloud of gas and pyroclasts interacts strongly with atmospheric gas which is entrained into the cloud⁷⁶ and heated by the pyroclasts⁷⁹. In some circumstances, a high, convecting eruption cloud can form while, in others, a 'collapsed' fountain is formed over the vent, feeding pyroclastic flows^{30,33,41} in the case of silicic magmas and lava flows in the case of mafic magmas. For any given exsolved volatile weight fraction, the eruption products will be prevented from forming a high, convecting cloud if the mass eruption rate of magma exceeds a critical value which depends on the volatile molecular weight and the atmospheric properties: the main reason is that geometric constraints prevent the cloud from entraining enough air to utilize enough of the heat content of the pyroclasts to create the required buoyancy^{33,41}. Table 1 gives the appropriate critical values for the volatile contents and molecular weights illustrated.

When a convecting eruption cloud can remain stable in a steady eruption (Fig. 2) its maximum height is proportional to the fourth root of the mass eruption rate^{39,40}; however, the constant of proportionality is a function of the physical properties of the atmosphere⁷⁸ and, as a result, eruption clouds would rise 5 times higher on Mars, and 3 times less high on Venus, than on Earth³⁹ for the same mass eruption rate. Thus, in general, pyroclastic material accumulating from eruption clouds is expected to be relatively more widespread on Mars and relatively less widespread on Venus than on Earth. One example of an air-fall ash deposit has so far been identified with some confidence on Mars³⁴, and other examples have been suggested⁸⁰.

Table 1 also shows that pyroclastic flow formation in silicic eruptions is likely to occur at lower mass eruption rates, and hence to be more common, on both Mars and Venus than on Earth. It has been suggested that some of the materials on the flanks of certain martian volcanoes are pyroclastic flow deposits⁸¹⁻⁸⁴, though often the evidence is indirect. There is as yet no direct information about possible pyroclastic deposits on Venus.

On planets with no appreciable atmosphere, such as the Moon and Io, equation (12) breaks down. Not far above the vent, the spatial separation of pyroclasts becomes large; the expanding gas loses good thermal contact with the pyroclasts and expands isentropically. The gas velocity approaches an asymptotic value^{45,47} as the gas density becomes very small and even small clasts decouple completely from the gas flow to follow ballistic trajectories^{25,46}. The ranges reached by the ejected clasts (and their temperatures on landing) are a very strong function of the size distribution of the clasts and the mass fraction, n , that they constitute of the total ejected material^{25,85}. The asymptotic ejection velocity, u_m , is given³⁸ with sufficient accuracy by

$$u_m = \left(\frac{2\gamma Q T_i n}{m(\gamma - 1)} \right)^{1/2} \quad (13)$$

where T_i is the temperature from which gas expansion begins and γ is the ratio of the specific heats of the gas.

For steady lunar basaltic eruptions in which the released gas was probably CO (ref. 28) present in amounts up to $n \approx 0.07$ wt% and in which magma was probably disrupted into largely sub-millimetre sized pyroclasts²⁵, $u_m \approx 40$ m s⁻¹ and pyroclastic blankets up to ~1 km in radius could have been formed. In non-steady, strombolian style explosive eruptions, in which low magma rise speed through the crust provides time for coalescence of gas bubbles, the gas fraction in the products of each explosion can be much larger than the average value in the magma as a whole and may have reached ~10 wt% (ref. 25) leading to $u_m \approx 500$ m s⁻¹ and ejecta ranges of many tens of kilometres.



Fig. 2 Mount St Helens eruption, 18 May 1980, looking north-east. This photograph, taken several hours after the eruption began, shows the convecting eruption cloud formed from explosively ejected gas and pyroclasts, and entrained atmospheric gas. In the background, steam and smoke rise from forest fires and the interaction of hot ash with rivers, lakes, snow and ice. US Geological Survey photograph 80-S3-137.

Eruptions on Io appear to involve either sulphur⁴⁷ or SO₂ (ref. 44) as the volatile phase, and eruption cloud heights up to 300 km imply $u_m \approx 1,000$ m s⁻¹. Such high velocities derived from these high molecular weight gases imply that most of the material released through the vent is in gaseous form^{47,85}, though condensation of the volatile back to the solid phase within the cloud is clearly important⁴⁴. The mean density of Io, and the preservation of hills and valleys with relief exceeding 2 km, argue for the crust and mantle being dominated by silicates^{69,86}; the observed sulphur volcanism must, therefore, almost certainly be driven by underlying silicate volcanism. However, the details of the mechanisms of heat and mass transfer across the boundary between the two systems have yet to be investigated⁸⁵.

Unsteady explosive eruptions

Other than in the case of eruptions on Io, the above discussion of explosive eruptions has assumed that the volatiles driving the explosions are magmatic, that is, they are present in the magma at the level in the crust at which the magma is last stored before eruption, and that the eruptions are steady in the sense that the discharge rates are slowly varying with time. But there are many exceptions to this pattern.

It is quite common on Earth for pyroclastic material to be produced by the expansion of magmatic volatiles and then erupted through water, or under circumstances where water has constant access to the vent. The result of the interaction is much greater comminution of the ejecta, and material is dispersed either steadily or episodically in phreato-plinian⁸⁷ or surtseyan⁸⁸ eruptions. A mechanism equivalent to the phreato-plinian style has been explored for Mars, where liquid water is not stable on the surface under the present atmospheric pressure but could be available via subsurface aquifers generated by heating of permafrost in volcanic intrusive episodes³⁴.

In low effusion rate eruptions of low viscosity magma on any planet, coalescence of small gas bubbles can lead to the intermittent (time scales of seconds to minutes) explosive emergence of large bubbles through the vent in the strombolian eruption



Fig. 3 Hadley Rille, a lunar sinuous rille located along the eastern edge of the Imbrium impact basin. The rille begins as an elongated trough at the junction between a major radial lineament and the main Imbrium basin scarp (Apennine Mountains) and then extends along the mountain front for about 55 km before turning north-west towards Mare Imbrium. Hadley rille is interpreted to be a lava channel widened by thermal erosion⁵⁵. Portion of Lunar Orbiter V frame 105M.

style^{25,89}. In low effusion rate eruptions of high viscosity magma a more important process is chilling of the magma which causes effusion to cease and allows time for the build up of gas pressure beneath a retaining carapace, either as a result of continuing exsolution of magmatic volatiles or as a result of heating of groundwater in surrounding country rocks^{38,90}; vulcanian explosions may result. A version of this mechanism (using only accumulating magmatic volatiles) has been invoked to explain the localized dark pyroclastic haloes surrounding small (1–2 km diameter), elongate explosion craters within the 120 km diameter impact crater Alphonsus on the Moon⁹¹.

Implications for planetary volcanism

These basic principles concerning the ascent and eruption of magma can be related to variables associated with differences in planetary size and environment. Predictions, and comparisons with observed volcanic landforms and deposits on the Earth, Moon, Mars, Venus and Io can then be made.

On Earth for basalts with viscosities exceeding 10^2 Pa s and negligible yield strengths, crustal fissure widths from 0.2 to 4 m would account for the full range of known mass eruption rates, including the $3 \times 10^4 \text{ kg s}^{-1} \text{ m}^{-1}$ estimated for the Columbia River flood basalt eruptions⁹². Significant thermal erosion leading to incision and widening of lava channels requires mass eruption rates much larger than commonly observed and is consequently not an important process on Earth. Gas exsolution from magma only occurs at relatively shallow depths (usually <2 km for terrestrial basalts), implying that the mass discharge rate of magma is mainly controlled by the rheological properties of the magma and the average width of the fissure system through which it rises. Strombolian style eruptions (intermittent explosive activity caused by bubble coalescence) will commonly occur in basaltic magmas if the rise speed in the crust is <0.5–1 m s^{-1} . If the rise speed exceeds this limit, relatively steady fire-fountaining will commonly take place in basaltic magmas. Fire-fountain height is directly related to the vertical velocity of the magma–gas mixture. Vertical velocity and fountain height

increase with increasing magma gas content and mass eruption rate and decrease with increasing magma viscosity. Where volatile content is high enough in silicic or mafic magmas, and magma is disrupted into small pyroclasts and gas, the ejected cloud will interact with the atmosphere to produce a convecting eruption cloud whose maximum height is proportional to the fourth root of the mass eruption rate. In many cases, cloud collapse occurs and pyroclastic flows are produced from silicic magmas.

On Venus the atmospheric pressure in the lowlands is so great (~90 bar) that magma cannot be disrupted at all by gas bubble growth unless the exsolved magma volatile content exceeds, for example, 2 wt% H_2O or 5 wt% CO_2 . The corresponding figures for the Venus highlands are 1 and 3 wt%, respectively. Such high volatile contents of relatively insoluble gases like CO_2 can be obtained only if the magma source region is at a depth ≥ 40 km. Thus, if magmas on Venus are very deficient in volatiles with high solubilities like that of H_2O , pyroclastic eruptions may never occur there^{93,94}. Detection of pyroclastic deposits or landforms on Venus could indicate that the present atmospheric characteristics have not prevailed throughout the history of Venus, or that volatile contents occur on Venus which are abnormally high by terrestrial standards. If pyroclastic eruptions do occur, eruption velocities will be smaller by a factor of 2–4 than those on Earth for the same volatile contents. The high atmospheric pressure will cause convecting cloud rise heights to be about three times lower than on Earth for the same mass eruption rate, and pyroclasts will be much less widely dispersed. Column collapse should be much more common and pyroclastic flow formation will be much more likely than on Earth if the ranges of mass eruption rate are similar. Pyroclastic flows will be much less mobile, however.

The high surface temperatures on Venus mean that surface cooling of lava flows is less efficient, which would favour a higher proportion of long single flows relative to comparable Earth eruptions. Higher upper crustal temperatures would inhibit magma cooling during ascent and favour correspondingly greater effusion rates.

On Mars, the low atmospheric pressure means that magma disruption into pyroclasts is much more likely to occur there than on Earth. The eruption velocities of gas and small pyroclasts on Mars will be higher than on Earth for the same volatile content by a factor of ~1.5. Martian eruption clouds will rise to heights greater than those on Earth by a factor of ~5 for the same mass eruption rate and pyroclasts will be much more widespread than on Earth, though relatively finer grained. Pyroclastic flow formation should be significantly more common on Mars than on Earth for similar ranges of eruption rate although recognition of such deposits is controversial^{95,96}.

Very long lava flows reported for some parts of Mars suggest that high mass eruption rates and relatively large crustal fissure widths have been common. Unless martian magmas are very volatile-poor, it should be more common on Mars than on Earth for effusive eruptions with low magma discharge rates to be accompanied by strombolian explosions, and the low martian atmospheric pressure and surface gravity will favour the production of low cinder cones which are several times wider than their equivalents would be on Earth. Quietly effusing martian lavas should contain gas vesicles up to four times larger than those found in terrestrial lavas.

On the Moon, the high mass eruption rates implied by the lengths of some lunar lava flows and sinuous rilles⁵⁷ (Fig. 3) require that lunar tectonic regimes produced fissure systems with widths up to ~10 m. Major impact basin ring systems and an extensional state of stress in the lithosphere, or an early thickening of the elastic lithosphere, may account for such large fissure widths. Low effusion rate eruptions could have produced cinder/spatter ridges up to 1 km wide along fissure sources. Central vents could have produced dome- or cone-like cinder deposits with diameters up to 5 km. High effusion rate eruptions should have produced optically thick eruption clouds from



Fig. 4 Dark-mantle deposits on the lunar Aristarchus Plateau. Expanding gas clouds from volcanic eruptions are believed to have accelerated fine-grained pyroclasts to great radial ranges to produce regionally significant mantling pyroclastic deposits¹⁰⁰. The source of the Aristarchus Plateau deposits may be the circular depression (Cobra Head) at the end of Schröters Valley. Later lava flows of higher albedo surround the plateau. The bright crater Aristarchus, approximately 40 km in diameter, is located in the lower right-hand corner of the plateau. Earth-based telescopic photograph from the Consolidated Lunar Atlas.

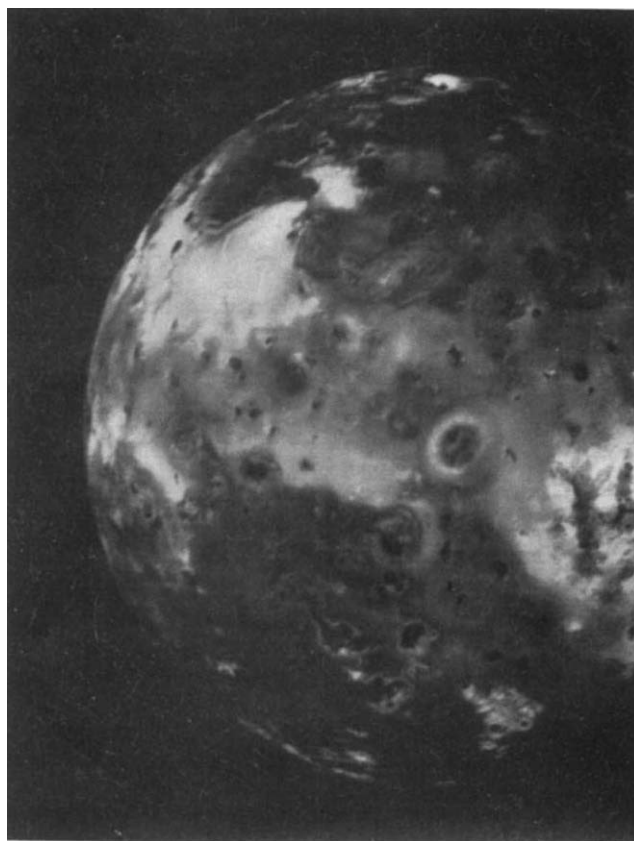


Fig. 5 Jupiter's satellite Io, showing a range of recent volcanic features including a doughnut-shaped eruption plume ~400 km in diameter marking the site of an active eruption. NASA Voyager photograph P-21209C.

which pyroclasts landed hot and coalesced to form lava flows. The landing zone could have been up to 4 km wide along elongate fissures and up to 6 km in diameter around central vents. The lava pool marking such a landing zone will have produced a thermally eroded depression feeding a sinuous rille if the eruption rate was $\geq 10^7 \text{ kg s}^{-1}$ (ref. 91). Numerous sinuous rilles are observed to have linear and circular depressions of the predicted sizes at their sources^{97,98}. Although lunar magma volatile contents were very low²⁸ by terrestrial standards, the absence of atmosphere should have ensured that magma disruption into pyroclasts was a common event. As the gas clouds expanded, pyroclasts became separated, gas density decreased, and particles decoupled and followed ballistic trajectories. Very widespread, thin deposits (up to 200 km in diameter) of sub-millimetre pyroclastic droplets should have been formed for a wide range of mass eruption rates in events where the bulk of the erupted material was retained near the vent in lava flows or cinder cones. The observed lunar dark mantle deposits^{99,100} are related to this style of activity (Fig. 4).

On Io the currently observed eruption clouds (Fig. 5) must contain large proportions (up to 50 wt%) of volatiles (apparently^{44,45,47} sulphur or SO_2) and may be the equivalent of terrestrial phreatomagmatic eruptions⁸⁵. If the mean size of the particles in the clouds is 100 μm or more then the particle motions must be largely ballistic^{43,46}. Cloud brightness modelling shows that parts of the clouds are optically thick which implies that eruption rates are $>10^7 \text{ kg s}^{-1}$ and that up to 40% of the entire heat flux from Io's interior may emerge through its active volcanoes⁸⁵. There is evidence that flows of liquid sulphur are erupted onto the surface of Io from many vents, but it is also possible that some of the flow features are of silicate lava. In either event, silicate liquids must be the main agents for the transfer of heat from the interior of Io through the lithosphere.

Outstanding problems

(1) Although the thermodynamics of eruptions on Io in which silicate magma evaporates sulphur compounds to produce explosive eruptions have been examined in depth⁴³⁻⁴⁷, the dynamical mixing of magma and sulphur has not been studied in detail. There should be many similarities between magma-

sulphur interactions on Io and magma- H_2O interactions on Earth—involving both the solid and liquid phases of the volatile in both cases. Although magma—liquid water eruptions are classifiable empirically on Earth^{87,88}, only rather general models for the process have been proposed¹⁰¹⁻¹⁰³ and a more general study taking account of geologically plausible magma-volatile interaction geometries is required.

(2) While the direct evidence (from lunar samples^{104,105} and from photogeological analyses²) for explosive lunar volcanism between 3,000 and 4,000 Myr ago is strong, there is still uncertainty as to the composition of the volatile phases. The Moon appears to be severely depleted in H_2O , and CO has been suggested^{28,106} as a major volatile; however, the possibility that sulphur compounds, halogens or volatile metals (like K or Na) may have contributed^{107,108} needs further exploration¹⁰⁹. This information would be generally relevant to the investigation of silicate volcanism on any body lacking an atmosphere (the Moon, Io or the larger asteroids).

(3) Most treatments of volcanism on Venus^{36,110} have assumed that magma volatiles reflect the dominance of CO_2 and the paucity of H_2O in the atmosphere^{111,112}. However, the similarities between the chemical compositions of the high-K rocks at the Venera 13 landing site⁷⁷ and rocks from the East African Rift system on Earth suggest (R. Macdonald, personal communication) the possibility that halogens may form a significant fraction of the exsolved magma volatiles on Venus. Little can be done to simulate explosive volcanic activity involving these gases on Venus (or the Earth), however, until better data are available on their solubilities in magmas. There is a general need for experimental measurements of the mutually dependent solubilities of H_2O , CO_2 , CO, SO_2 , H_2S and the halogens in a wide range of magma types.

(4) Ideas on the range of magma types to be considered as candidates for eruption on a given planet are strongly influenced

by the range of magma types observed on the Earth, the range of types sampled on the Moon and the morphological features of the large volcanoes photographed on Mars. Comparison of these data sets suggests that basaltic volcanism is the major common denominator. However, appreciable quantities of highly silicic magma are produced on Earth by the chemical evolution of basaltic magma bodies stored in shallow crustal reservoirs, and this process is not itself controlled by the plate-tectonic environment in which it takes place. There is no *a priori* reason why silicic magmas should not be produced on the other silicate planets⁵² and, indeed, there do exist photographed but unsampled lunar surface features which may represent the extrusion of silicic lavas¹¹³. The eruption of volatile-bearing silicic magmas on Earth is commonly very explosive and all explosive volcanic processes are more extreme on low atmospheric pressure planets²⁵. Consequently, the products of silicic explosive eruptions on the Moon, Mars, Mercury or Io could be so finely-grained and widely dispersed as to produce no positive construct around the vent and no well-defined ejecta deposit. They would, therefore, be very hard to detect in currently available data sets.

(5) Most models of planetary interior dynamics have been

developed specifically for the Earth. There exists no general model, however, for the interaction of a relatively rigid planetary lithosphere with the less viscous planetary interior. The need for such a general model was not recognized while the Earth was the only planet known to have a dynamically active mantle and a tectonically active lithosphere. However, recent planetary geology investigations have shown that, despite the high level of localized volcanic activity on Io (where the relative heat flux from the interior is much greater than on Earth) plate tectonic activity appears to be entirely absent from the planet. It seems clear that these contrasting styles of heat transfer will only be understood, and that the current debate about the presence or absence of plate tectonic activity on Venus will only be resolved, in terms of a general model which relates the ways in which a planet can lose heat (by volcanic advection or direct conduction through the lithosphere) to the thickness and stability of the lithosphere.

We thank the William F. Marlar Memorial Foundation and the NASA Planetary Geology Program for financial support; and Peter Mouginiis-Mark, Geoff Wadge and Jonathan Fink for helpful comments; and Jackie Dixon and Susan Sharpton for help in preparation of the manuscript.

1. Basaltic volcanism study project *Basaltic Volcanism on the Terrestrial Planets* (Pergamon, New York, 1981).
2. Head, J. W. *Rev. Geophys. Space Phys.* **14**, 265–300 (1976).
3. Carr, M. H. *J. geophys. Res.* **78**, 4049–4062 (1973).
4. Greeley, R. & Spudis, P. *Rev. Geophys. Space Phys.* **19**, 13–41 (1981).
5. Strom, R. G., Trask, N. J. & Guest, J. E. *J. geophys. Res.* **80**, 2478–2507 (1976).
6. Phillips, R. J. & Malin, M. C. in *Venus* (eds Hunt, D. M., Colin, L. & Donahue, T. M.) (University of Arizona Press, Tucson, in the press).
7. Solomon, S. C. & Head, J. W. *J. geophys. Res.* **87**, 9236–9246 (1982).
8. Ramberg, H. *Deformation and the Earth's Crust as Studied by Centrifugal Models* (Academic, New York, 1967).
9. Weertman, J. *J. geophys. Res.* **76**, 8544–8553 (1971).
10. Elder, J. *The Bowls of the Earth* (Oxford University Press, 1976).
11. Fedotov, S. A. *Int. geol. Rev.* **6**, 671–680 (1977).
12. Marsh, B. D. *Phil. Trans. R. Soc. A288*, 611–625 (1978).
13. Marsh, B. D. & Kantha, L. H. *Earth planet. Sci. Lett.* **39**, 435–443 (1978).
14. Marsh, B. D. *Am. J. Sci.* **282**, 808–855 (1982).
15. Weertman, J. *J. geophys. Res.* **76**, 1171–1183 (1971).
16. Johnson, A. M. & Pollard, D. D. *Tectonophysics* **18**, 261–309 (1973).
17. Pollard, D. D. & Muller, O. H. *J. geophys. Res.* **81**, 975–984 (1976).
18. Shaw, H. R. in *Physics of Magmatic Processes* (ed. Hargraves, R. B.) (Princeton University Press, 1980).
19. Spera, F. J. in *Physics of Magmatic Processes* (ed. Hargraves, R. B.) (Princeton University Press, 1980).
20. Fedotov, S. A. *Int. geol. Rev.* **20**, 33–48 (1978).
21. Hardee, M. C. & Larson, D. W. *J. Volcanol. geotherm. Res.* **2**, 299–308 (1977).
22. Delaney, P. T. & Pollard, D. D. *U.S. geol. Surv. Prof. Pap.* **1202**, 1–61 (1981).
23. Solomon, S. C. *Proc. 6th Lunar planet. Sci. Conf.* 1021–1042 (1975).
24. Davies, P. A. & Stephenson, A. *Phys. Earth planet. Int.* **14**, P13–16 (1977).
25. Wilson, L. & Head, J. W. *J. geophys. Res.* **86**, 2971–3001 (1981).
26. Shaw, H. R. & Swanson, D. A. *Proc. 2nd Columbia River Basalts Symp.* (E. Washington State College Press, Cheney, 1970).
27. McGetchin, T. R. & Ullrich, W. G. *J. geophys. Res.* **78**, 1833–1853 (1973).
28. Housley, R. M. *Proc. 9th Lunar planet. Sci. Conf.* 1473–1484 (1978).
29. Pai, S. I., Hsu, Y. & O'Keefe, J. A. *Proc. 9th Lunar planet. Sci. Conf.* 1485–1508 (1978).
30. Wilson, L., Sparks, R. S. J. & Walker, G. P. L. *Geophys. J. R. astr. Soc.* **63**, 117–148 (1980).
31. Kieffer, S. W. *J. geophys. Res.* **82**, 2895–2904 (1977).
32. Kieffer, S. W. & Delaney, J. M. *J. geophys. Res.* **84**, 1611–1620 (1979).
33. Wilson, L. & Head, J. W. *Lunar planet. Sci.* **12**, 1194–1196 (1981).
34. Mouginiis-Mark, P. J., Wilson, L. & Head, J. W. *J. geophys. Res.* **87**, 9890–9904 (1982).
35. Head, J. W., Wilson, L. & Mouginiis-Mark, P. J. *Int. Conf. on Venus Environment*, Palo Alto, California (1981).
36. Head, J. W. & Wilson, L. *Lunar planet. Sci.* **13**, 312–313 (1982).
37. Bennett, F. D. *Am. J. Sci.* **274**, 648–661 (1974).
38. Wilson, L. *J. Volcanol. Geotherm. Res.* **8**, 297–313 (1980).
39. Wilson, L., Sparks, R. S. J., Huang, T. C. & Watkins, N. D. *J. geophys. Res.* **83**, 1829–1836 (1978).
40. Settle, M. J. *J. Volcanol. Geotherm. Res.* **3**, 309–324 (1978).
41. Sparks, R. S. J., Wilson, L. & Hulme, G. *J. geophys. Res.* **83**, 1727–1739 (1978).
42. Morabito, L. A., Synnot, S. P., Kuperman, P. N. & Collins, S. A. *Science* **204**, 972–976 (1979).
43. Cook, A. F., Shoemaker, E. M. & Smith, B. A. *Nature* **280**, 743–746 (1979).
44. Smith, B. A., Shoemaker, E. M., Kieffer, S. W. & Cook, A. F. *Nature* **280**, 738–743 (1979).
45. Kieffer, S. W. in *Satellites of Jupiter* (ed. Morrison, D.) (University of Arizona Press, Tucson, 1982).
46. Strom, R. G., Schneider, N. M., Terrile, R. J., Cook, A. F. & Hansen, C. *J. geophys. Res.* **86**, 8593–8620 (1982).
47. Reynolds, R. T., Peale, S. J. & Cassen, P. *Icarus* (in the press).
48. Peale, S. J., Cassen, P. & Reynolds, R. T. *Science* **203**, 892–894 (1979).
49. Blake, S. *Nature* **289**, 783–785 (1981).
50. Hulme, G. *Geophys. J. R. astr. Soc.* **39**, 361–383 (1974).
51. Moore, H. J., Arthur, D. W. G. & Schaber, G. G. *Proc. 9th Lunar planet. Sci. Conf.* 3351–3378 (1978).
52. Rutherford, M. J. & Hess, P. C. *Lunar planet. Sci.* **12**, 915–917 (1981).
53. Fink, J. *Geology* **8**, 250–254 (1981).
54. Plescia, J. B. *Icarus* **45**, 586–601 (1981).
55. Hulme, G. *Mod. Geol.* **4**, 107–117 (1973).
56. Hulme, G. *Geophys. Surv.* **5**, 245–279 (1982).
57. Hulme, G. & Fielder, G. *Phil. Trans. R. Soc. A285*, 227–234 (1977).
58. Walker, G. P. L. *Phil. Trans. R. Soc. A274*, 107–118 (1973).
59. Head, J. W. & Wilson, L. *Lunar planet. Sci.* **12**, 427–429 (1981).
60. Whitford-Stark, J. L. *Earth Sci. Rev.* **18**, 109–168 (1982).
61. Mouginiis-Mark, P. J., Zisk, S. H. & Downs, G. S. *Nature* **297**, 546–550 (1982).
62. Carr, M. H., Greeley, R., Blasius, K. R., Guest, J. E. & Murray, J. B. *J. geophys. Res.* **82**, 3985–4015 (1977).
63. Schaber, G. G., Horstman, K. C. & Dial, A. L. *Proc. 9th Lunar planet. Sci. Conf.* 3433–3458 (1978).
64. Sparks, R. S. J. *J. Volcanol. geotherm. Res.* **3**, 1–37 (1978).
65. Sparks, R. S. J., Pinkerton, H. & Hulme, G. *Geology* **4**, 269–271 (1976).
66. Pinkerton, H. & Sparks, R. S. J. *J. Volcanol. geotherm. Res.* **1**, 167–182 (1976).
67. Wadge, G. *Geology* **6**, 503–506 (1978).
68. Malin, M. C. *Geology* **8**, 306–308 (1980).
69. Sagan, C. *Nature* **280**, 750–753 (1979).
70. Greeley, R., Fink, J. & Park, S. *Trans. Am. geophys. Un.* **62**, 1080 (1981).
71. Baloga, S. M., Pieri, D. C., Sagan, C. & Nelson, R. M. *Trans. Am. geophys. Un.* **62**, 1080 (1981).
72. Landau, L. D. & Lifshitz, E. M. *Fluid Mechanics* (Pergamon, New York, 1959).
73. Clough, B. J., Wright, J. V. & Walker, G. P. L. *Nature* **289**, 49–50 (1981).
74. Head, J. W. & Wilson, L. *Volcanic Processes on Venus* (in preparation).
75. Mutch, T. A., Arvidson, R. E., Jones, K. L., Head, J. W. & Saunders, R. S. *The Geology of Mars* (Princeton University Press, 1976).
76. Carr, M. M. *The Surface of Mars* (Yale University Press, New Haven, 1981).
77. Barsukov, V. L. *Pap. at 13th lunar planet. Sci. Conf.* Houston, Texas (1982).
78. Morton, B. R., Taylor, G. & Turner, J. S. *Proc. R. Soc. A234*, 1–23 (1956).
79. Wilson, L. *Geophys. J. R. astr. Soc.* **45**, 543–556 (1976).
80. Ward, A. W. *J. geophys. Res.* **84**, 8147–8166 (1979).
81. Reimers, C. E. & Komar, P. D. *Icarus* **39**, 88–110 (1979).
82. Scott, D. M. & Tanaka, K. L. *Rep. planet. Geol. Prog. 1980 NASA-TM 82385*, 255–257 (1980).
83. Morris, E. C. *Rep. planet. Geol. Prog. 1980 NASA-TM 82385*, 252–254 (1980).
84. Petersen, C. M. *Lunar planet. Sci.* **12**, 828–829 (1981).
85. Wilson, L. & Head, J. W. *Lunar planet. Sci.* **12**, 1191–1193 (1981).
86. Carr, M. H., Masursky, H., Strom, R. G. & Terrile, R. J. *Nature* **280**, 729–733 (1979).
87. Self, S. & Sparks, R. S. J. *Bull. Volcan.* **41**, 1–17 (1978).
88. Walker, G. P. L. *Geol. Rdsh.* **62**, 431–446 (1973).
89. Blackburn, E. A., Wilson, L. & Sparks, R. S. J. *J. geol. Soc. Lond.* **132**, 429–440 (1976).
90. Self, S., Wilson, L. & Nairn, I. A. *Nature* **277**, 440–443 (1979).
91. Head, J. W. & Wilson, L. *Proc. 10th Lunar planet. Sci. Conf.* 2861–2897 (1979).
92. Swanson, D., Wright, T. L. & Helz, R. T. *Am. J. Sci.* **275**, 877–905 (1975).
93. Garvin, J. B., Head, J. W. & Wilson, L. *Lunar planet. Sci.* **13**, 253–254 (1982).
94. Head, J. W. & Wilson, L. *Lunar planet. Sci.* **13**, 312–313 (1982).
95. Scott, D. H. & Tanaka, K. L. *J. geophys. Res.* **87**, 1179–1190 (1982).
96. Schultz, P. H. & Lutz-Garihan, A. B. *3rd int. Colloq. on Mars, Lunar and Planetary Institute Contr. No.* 441, 229–231 (1981).
97. Greeley, R. *Moon* **3**, 289–314 (1971).
98. Strain, P. & El-Baz, F. *Moon* **16**, 221–229 (1976).
99. Lucchitta, B. K. *Proc. 4th lunar Sci. Conf.* 149–162 (1973).
100. Head, J. W. *Proc. 5th lunar Sci. Conf.* 207–222 (1974).
101. Colgate, S. A. & Sigurgeirsson, T. *Nature* **244**, 552–555 (1973).
102. Schmincke, H.-U. *Geol. Jahrb.* **A39**, 3–45 (1977).
103. Peckover, R. S., Buchanan, D. J. & Ashby, D. E. *Nature* **245**, 307 (1973).
104. Heiken, G. & McKay, D. S. *Proc. 5th lunar Sci. Conf.* 843–860 (1974).
105. Heiken, G. & McKay, D. S. *Proc. 8th lunar Sci. Conf.* 3243–3255 (1977).
106. Sato, M. *Proc. 7th lunar Sci. Conf.* 1323–1344 (1976).
107. Chou, C. L., Boynton, W. V., Sundberg, L. L. & Wasson, J. T. *Proc. 6th lunar Sci. Conf.* 1701–1727 (1975).
108. Meyer, C., McKay, D. S., Anderson, D. H. & Butler, P. *Proc. 6th lunar Sci. Conf.* 1673–1699 (1975).
109. Cirline, E. M., Housley, R. M. & Grant, D. W. *Proc. 9th lunar planet. Sci. Conf.* 2049–2063 (1978).
110. Garvin, J. B., Head, J. W. & Wilson, L. *Icarus* **52**, 365–372 (1982).
111. Ahrens, T. J. *Adv. Space Res.* **1**, 177–187 (1981).
112. Goettel, K. A., Shields, J. A. & Decker, D. A. *Proc. 12th Lunar planet. Sci. Conf.* 1507–1516 (1981).
113. Head, J. W. & McCord, T. B. *Science* **199**, 1433–1436 (1978).

Molecular analysis of a cell lineage

Kim Nasmyth

Laboratory of Molecular Biology, MRC Centre, Hills Road, Cambridge CB2 2QH, UK

Mating type interconversion has a precise lineage which serves to minimize the time taken for yeast cells to achieve the diploid state. The HO gene either encodes or regulates an endonuclease which initiates the interconversion process. Expression of this gene is switched on during the G₁ phase of mother cells and not at all during the cell cycle of their daughters. This behaviour can explain what is known about the lineage.

MATING type in yeast is determined by the *MAT* locus which has two alleles, *a* and α . The predominant vegetative phase is a non-mating diploid cell which is heterozygous for *MAT* (that is, *MATa/MAT α*). Starvation of this cell leads, through meiosis and sporulation, to the production of a tetrad of haploid spores, two having an *a* and two an α mating type. In heterothallic (*ho*) strains, *MATa* and *MAT α* are stable mendelian alleles. Germination of individual heterothallic spores thus gives rise to vegetative haploid cells of a given homogeneous mating type. The diploid phase of the life cycle can only be regained by conjugation between cells of opposite mating type, that is, of different clonal origin¹.

Most wild yeast isolates, in fact, have homothallic life cycles

(P. Philippsen, personal communication), which means that single haploid spores of a given mating type can give rise to non-mating *a/a* diploid cells without encountering cells of a different mating type, as is obligatory for heterothallic strains. Hawthorne² proposed that the homothallism gene *HO* (which is allelic to *ho*) acts by directing the efficient mutation of one mating type to another, giving rise to colonies of heterogeneous mating type, in which diploidization is achieved, just as in heterothallic strains, by conjugation between cells of opposite mating types (see Fig. 1A). Both genetic (reviewed by Herskowitz and Oshima³) and physical (reviewed by Nasmyth⁴) analyses have confirmed this hypothesis. It is now clear that interconversion between the *a* and α alleles of the *MAT* locus

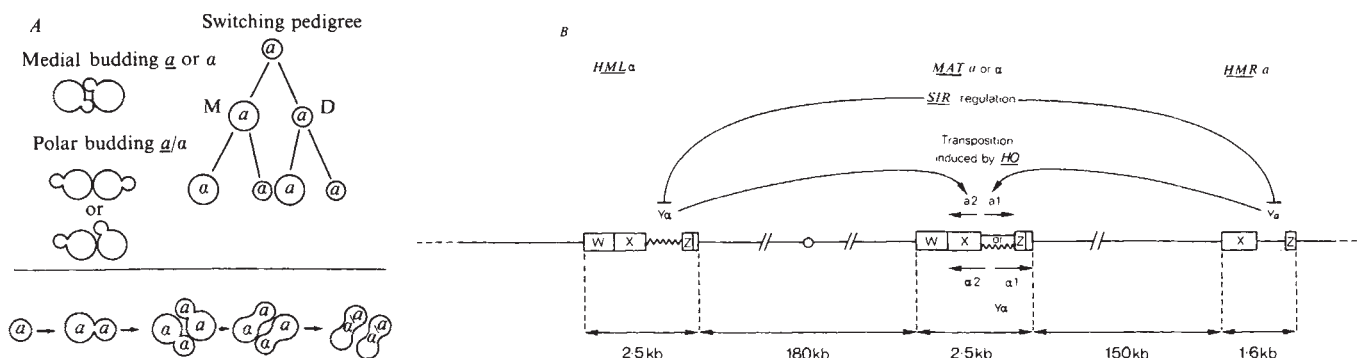


Fig. 1 Diploidization by mating type switching in homothallic yeast. **A**, Budding patterns, the pedigree of switching and their involvement in diploidization. **B**, The physical structure of yeast mating type loci on chromosome 3. **A**, Homothallic spores never switch mating type during their first cell cycle. The mother cell produced by this first division (M) invariably ($\geq 80\%$) switches during the next cell cycle to produce mother and daughter cells of changed mating type (here, *a* $\rightarrow$ α). In contrast, the daughter cell produced by the first division never switches during its next cell cycle. The mother cell produced by this division, however, will switch during its next cell cycle. Only mother cells, therefore, are capable of switching⁶. The question is what is the function of this highly regulated switching behaviour? A strong case can be made that it serves to accelerate the diploidization process. In order to see this, it is necessary to consider the budding pattern of yeast. Cells which are homozygous for *MAT* (that is, *a* or α haploids) generally produce their next bud at the junction between mother and daughter cell—this is called medial budding¹⁹. In contrast, cells which are heterozygous for *MAT* (*a/a* diploids) generally produce their next bud either at the opposite pole from the previous bud or at random with respect to it—this is called polar budding. While the mating type control over budding pattern and the switching pedigree individually appear enigmatic, both make more sense if considered together as a coordinated process for accelerating diploidization (see diagram below the line). Medial budding and non-switching of daughter cells together ensure that each α cell is opposite an *a* partner at the four-cell stage⁴. **B**, Mating type is determined by the *MAT* locus, which is situated on the right arm of chromosome 3. The two alleles of *MAT* (*MATa* and *MAT α*) differ in their DNA sequences. *MATa* contains a 642-bp *Ya* sequence which is replaced at *MAT α* by a non-homologous 747-bp *Y α* sequence. As a result of these different Y sequences alone, *MATa* and *MAT α* each produce two unique transcripts (which determine the *a* or α mating type). A silent copy of *Ya* exists at *HML α* , which is near the left telomere, and a silent copy of *Y α* exists at *HMR α* , which is near the right telomere of the same chromosome (3). *HML α* and *HMR α* are kept silent by four *SIR* genes. Mutation of any one of *SIR* 1, 2, 3 or 4 allows both *HML α* and *HMR α* to produce the mating-type transcripts which are normally only expressed at *MAT*. During a mating type interconversion, *HML α* and *HMR α* serve as donors of *Ya* or *Y α* information to *MAT*. For example, when an *a* cell switches to α , the *Ya* sequence residing at *MATa* is replaced by the *Y α* sequence residing at *HML α* . Unlike the *MAT* DNA, the silent locus is not altered by the interconversion, that is, it is not an exchange but a unidirectional flow of information from *HML* or *HMR* to *MAT*. A sequence called X (on the left) and one called Z1 (on the right) surround the Y sequence at *HML*, *HMR* and *MAT*. *MAT* and *HML* are also homologous for the W sequence (to the left of X) and Z2 (to the right of Z1). Although *HML* and *HMR* usually contain *Ya* and *Y α* , respectively, in some strains they contain a Y sequence of the opposite type. *HML* contains W and Z2 and *HMR* does not, irrespective of whether they contain *Ya* or *Y α* .

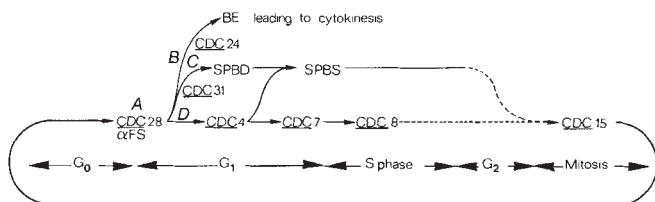


Fig. 2 A map of certain events during the yeast cell cycle, concentrating on events occurring during the G_1 phase. As cells grow out from G_0 , the first event which they traverse has been called $START^9$. This is the step at which α cells are arrested by α factor (marked as αFS) and which requires the function of the *CDC28* gene product (as well as *CDC36*, 37 and 39). Three independent pathways diverge thereafter: one leading to bud emergence (BE) requiring *CDC24*, one leading to duplication of the spindle pole body (SPBD) requiring *CDC31*, and one leading to DNA replication requiring *CDC4* and *CDC7*. *cdc4* mutants not only fail to initiate DNA replication but also fail to undergo spindle pole body separation (SPBS). *CDC8* is required for chain elongation during S phase and *CDC15* is required for a late stage in mitosis. Although it varies rather between strains, the mean generation time of yeast is ~150 min at 23 °C and ~90 min at 37 °C (distances along the abscissa are not proportional to time).

is accomplished by the replacement of α - or α -specific DNA sequences at *MAT* (each are unique and about 700 base pairs (bp) long) by copies of the opposite type which reside at different silent loci: *HML* (for α) and *HMR* (for a). Figure 1B summarizes this process.

Hicks and Herskowitz⁵ have shown that the interconversion of mating types in homothallic cells has a precise lineage which is analogous to the generation of new cell types by stem cells (Fig. 1A). It is with two aspects of this lineage⁶ that I am concerned here: (1) both progeny of a cell which has changed mating type always have the same new mating type, and (2) only cells which have previously given rise to a daughter cell, that is, mother cells, are capable of switching (spores are honorary daughters). The first aspect suggests that switching occurs at a specific stage in the cell cycle, that is, before *MAT* replication, and the second aspect that there exists a specific mechanism for the asymmetric segregation (between mother and daughter) of the potential to switch. This temporal and spatial regulation of switching serves to accelerate the diploidization process (see Fig. 1A).

There is now considerable evidence that mating type interconversion is initiated by a double strand cleavage of a specific sequence at the *MAT* locus^{4,7}. The target of this cleavage is a

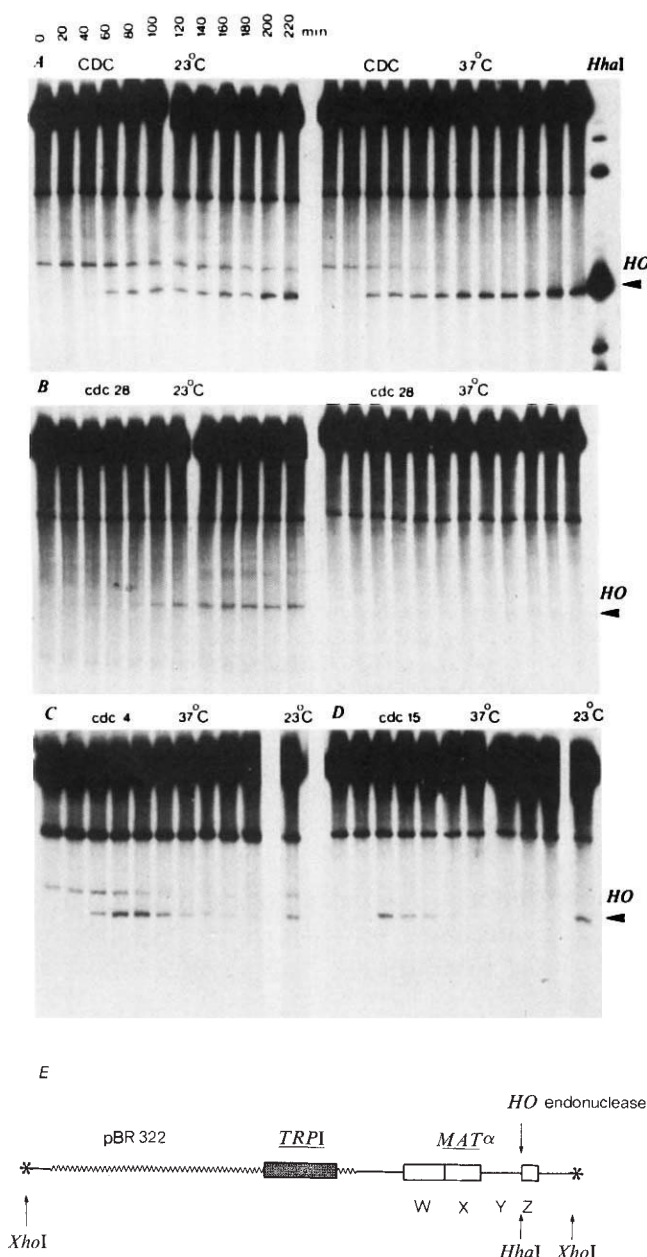


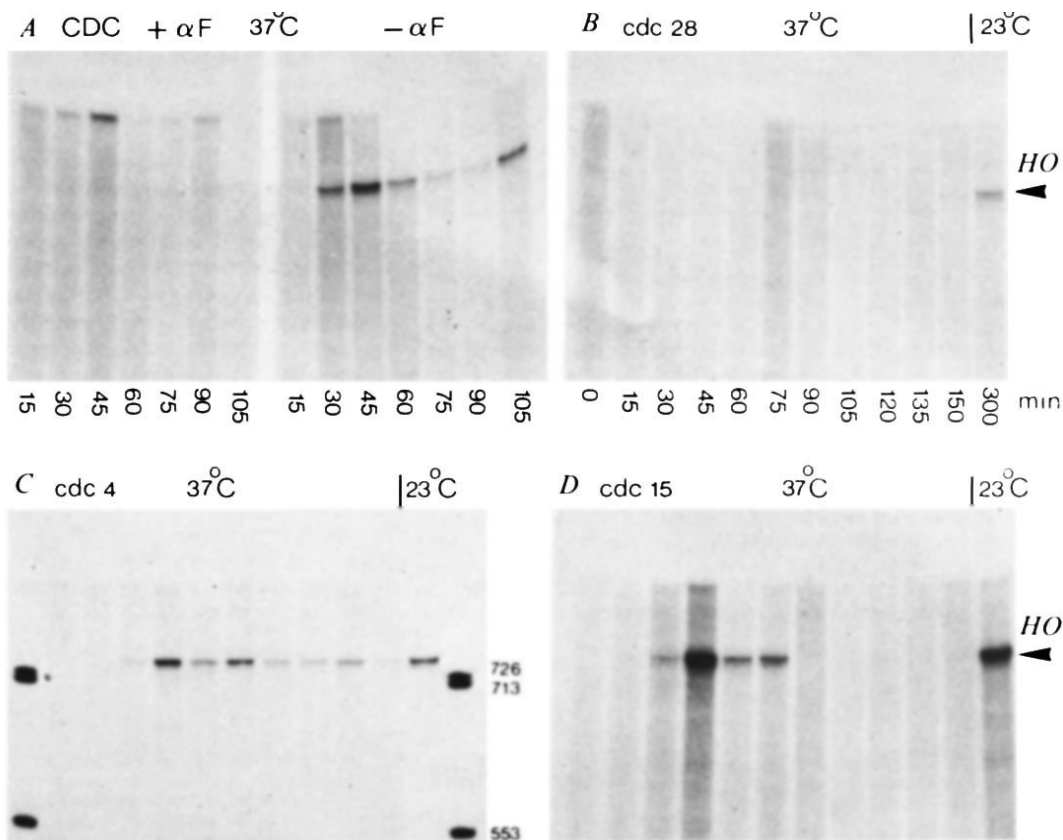
Fig. 3 Level of *HO* endonuclease in various *cdc* mutants on outgrowth from G_0 . A, Wild-type (strain S1) at 23 °C (left) and 37 °C (right); B, *cdc28-4* (S2) at 23 °C (left) and 37 °C (right); C, *cdc4-1* (S3) at 37 °C; D, *cdc15-2* (S4) at 37 °C; E, the end-labelled substrate. An arrow marks the band produced by *HO*-dependent cutting at the Y-Z junction. The lane marked *HhaI* in A shows the substrate digested with *HhaI* instead of a yeast extract. The lanes marked 23 °C in C and D represent activity in cells incubated at 23 °C for 300 min. Up to four bands were revealed by autoradiography. The top and most prominent corresponds to undigested substrate, the second is due to an endonuclease present at a similar level in all extracts, the third is strain-specific and more prominent in stationary phase extracts, and the fourth and bottom band is due to the *HO*-dependent endonuclease; its break coincides with a *HhaI* site at the Y-Z junction (see A).

Methods: G_0 -arrested cultures were prepared by growing cells on YEED plates at 25 °C until confluent (0.1 ml of a stationary phase culture was spread evenly on a plate and incubated at 25 °C for 48–60 h). Over 95% of these cells are unbudded cells with a single nucleus. To inoculate, the cells were suspended in 5 ml of fresh YEED added to each plate, and 1.5 ml of the suspension added to 100 ml of fresh medium at 23 °C or 37 °C. Note that these cultures are not very synchronous in their cell cycle progression. Extracts were prepared from 7.5 ml of cells. Cells were cooled on ice for 5 min, collected by centrifugation, washed once with 1 ml of ice-cold extraction buffer (10% glycerol, 100 mM Tris-HCl pH 7.7, 1 mM EDTA, 1 mM dithiothreitol (DTT), 0.5 mM phenylmethylsulphonyl fluoride, PMSF), and suspended in 40 μ l of extraction buffer. Cells were then vortexed with glass beads (Ballotini no. 11) at 4 °C for 2.5 min, centrifuged for 3 min in a microfuge at 4 °C, and 18 μ l of supernatant removed and frozen at -70 °C. The endonuclease substrate was prepared by digesting DNA of plasmid YRp7. *MAT* α .87 (ref. 20) with *XhoI*, followed by end-labelling using T4 polymerase²¹. About 15 μ Ci of α -labelled dATP was incorporated into the ends of 1.1 μ g of linearized plasmid. Assays were performed by adding 2 μ l of extract to 15 μ l of 33 mM Tris-OAc pH 8.0, 0.12 M KOAc, 12 mM Mg(OAc)₂, 100 μ g ml⁻¹ bovine serum albumin 1 mM ATP, 0.5 mM DTT, containing 16.5 ng of substrate. Incubation was at 30 °C for 90 min. The reaction was stopped by adding 5 μ l of 4% SDS, 0.05 M Tris-HCl pH 8.0, 0.04 M EDTA, 250 μ g ml⁻¹ proteinase K, followed by incubation at 37 °C for a further 20 min. The mixture was then extracted once with an equal volume of phenol-chloroform, 18 μ l of the aqueous phase were added to 2.5 μ l of 30% Ficoll, 50 mM EDTA, 0.25% bromophenol blue, 0.25% xylene cyanol, and electrophoresed on a 20 cm, 4% (30:1) acrylamide gel in 0.5 $\times$ TBE for 2–3 h at 220 V. Gels were dried before autoradiography. Overnight exposures without an intensifying screen were generally sufficient. Experiments in which active extracts (from mother cells) were serially diluted with inactive ones (from equivalent time points of daughter cells) showed that, although the activities in this assay are not strictly proportional to the amount of extract added, the assay does give a quantitative measure of the specific activity of the *HO* endonuclease. The extracts shown were prepared at 20-min intervals after inoculation. All strains used were of the genotype *HML* α , *MAT* α , *HMR* α , *HO*. The lack of any α DNA in these strains prevented them from becoming a/α diploids (in which *HO* is repressed).

Fig. 4 Level of *HO* RNA in various *cdc* mutants on outgrowth from G_0 . **A**, Wild type (strain S1) at 37 °C in the presence (left) and absence (right) of 40 $\mu\text{g ml}^{-1}$ α factor (αF ; Sigma); **B**, *cdc 28-4* (S2) at 37 °C; **C**, *cdc 4-1* (S3) at 37 °C; **D**, *cdc 15-2* (S4) at 37 °C. Arrowhead marks the band due to *HO* RNA. The lanes marked 23 °C are RNA from cells incubated at 23 °C for 300 min. Times after inoculation into fresh medium are shown in **A** and **B**; the times for **C** and **D** were as shown in **B**.

Methods: The preparation of G_0 cells and their inoculation into fresh medium were as described in Fig. 3 legend. RNA samples were prepared from 7.5 ml of cells. Cells were cooled on ice for 5 min, collected by centrifugation, washed once with 1 ml of ice-cold RNA buffer (0.05 M Tris-HCl pH 8.0, 0.1 M NaCl, 0.01 M EDTA) and the

pellet of cells frozen in microfuge tubes and stored at -70 °C. To prepare RNA, the cells were resuspended in 40 μl of 0 °C RNA buffer, and vortexed with glass beads at 4 °C for 2.5 min; 40 μl of fresh RNA buffer were added to the broken cells together with 3 μl of 25% SDS and 100 μl of phenol saturated with 0.1 M Tris-HCl pH 7.5 (with mixing in between each step). The samples were centrifuged at 10,000g for 2 min, and 70 μl of the aqueous phase removed. The samples were extracted once more with phenol and once with a 24:1 mixture of CHCl_3 -isoamyl alcohol before nucleic acids were ethanol-precipitated. The precipitate was washed once with 95% ethanol and dissolved in 50 μl of TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA). Hybridization probes were prepared from recombinant clones of the M13 vector mp9 (ref. 22) containing the 0.87-kb *Bam*HI-*Hind*III fragment from the *HO* gene¹⁰. Using single-stranded phage DNA as template, Klenow polymerase was used to extend a 17-nucleotide long M13 primer²³ across the insert in the presence of 200 Ci mmol^{-1} α dATP. The resulting double-stranded DNA was cleaved with *Hind*III, and the radioactive strand of the 870-nucleotide fragment was isolated by electrophoresis through a 7 M urea-4% polyacrylamide gel. To measure the relative level of *HO* RNA, 12 μg of RNA were co-precipitated with an excess of hybridization probe, redissolved and hybridized at 65 °C, treated with S_1 endonuclease, and electrophoresed through a 7 M urea-4% polyacrylamide gel as described previously²⁴. By this method, S_1 treatment of *HO* RNA/DNA hybrids yielded a ~750-nucleotide long fragment (implying that the 5' end of the *HO* transcript is 750 nucleotides from the *Bam*HI site). The amount of this fragment was directly proportional to the amount of RNA used (results not shown). A similar experiment with an mp8 clone containing the *Bam*HI-*Hind*III fragment gave no such protection.



10-bp recognition sequence situated at the Y-Z junction (Fig. 1B). It has been reported recently (R. Kostriken and F. Heffron, personal communication) that the *HO* gene encodes or regulates an endonuclease of the appropriate specificity. As the recognition sequence for this enzyme lies not only at the recipient *MAT* locus but also at the two donor loci *HML* and *HMR*, and yet only cuts the former *in vivo*, it follows that the sequence can be masked. It seems likely that this masking is due to the sequence being embedded in a different chromatin structure at *HML* and *HMR* from that at *MAT*⁸. Temporal regulation of mating type interconversion could therefore be accomplished by at least two mechanisms: (1) variation in the susceptibility of the *MAT* recognition sequence to cleavage by the endonuclease and (2) variation of the level of endonuclease activity. This paper tests the latter hypothesis.

Cell cycle analysis

The first question is when in the cell cycle does endonuclease appear and disappear? To answer this, I have used cell cycle mutants rather than analyse synchronous cultures. I have thus asked whether or not the appearance of endonuclease depends on particular cell cycle events. In practice, this has been performed by arresting strains harbouring temperature-sensitive (*ts*) alleles of various *CDC* genes⁹ in G_0 by starvation at the permiss-

ive temperature (23 °C). Re-inoculation of these starved G_0 cultures into fresh medium at the restrictive temperature (37 °C) results in the cells' traversing the cell cycle until they reach the block imposed by the *ts cdc* allele. In these circumstances, some *cdc* mutants will traverse almost a whole cell cycle—for example, *cdc 15* cells will traverse G_1 , S phase and G_2 before arresting at a late stage in mitosis. Other *cdc* mutants, for example, *cdc 28*, will arrest before they have even begun the G_1 phase. Figure 2 describes certain relevant features of the yeast cell cycle and the phenotypes of various *cdc* mutations.

HO endonuclease activity during outgrowth of *CDC*⁺ and *cdc*⁻ mutants

Double strand endonuclease activity has been assayed by incubating end-labelled naked *MAT* α DNA with crude extracts, followed by gel electrophoresis on native polyacrylamide gels (see Fig. 3). Several site-specific endonuclease activities are detected using this assay. An activity cleaving at the Y-Z junction is easily recognized as it coincides with a *Hha*I site at *MAT* α . This activity is found neither in *a*/ α cells where the *HO* gene is repressed¹⁰ nor in *ho* strains (results not shown). It will henceforth be called the *HO* endonuclease.

Figure 3 compares the temporal profile of *HO* endonuclease after various G_0 -arrested *cdc* mutants were inoculated into

Table 1 *cdc* Mutants used in the analysis

Strain	Name	Genotype	Source
S1	DC113	<i>HMLa HMRa MATa HO his5 ade5 met4 ura4</i>	Ref. 7
S2	K554	<i>HMLa HMRa MATa HO cdc28-4 met-4 or met2</i>	Present work
S3	K556	<i>HMLa HMRa MATa HO cdc4-1 met-4</i>	Present work
S4	K557	<i>HMLa HMRa MATa HO cdc15-2 met-4</i>	Present work
S5	PB31	<i>HMLa HMRa MATa ho cdc31-1 ade2 or 1 leu2 trp1 lys1</i>	J. Yocum

fresh medium at 23 °C or 37 °C. No activity was detected in the starting unbudded G_0 cultures. In the wild type (Fig. 3A), activity appears by 40 min at 37 °C and by 60 min at 23 °C. Thereafter, it slowly rises with time. The onset of activity just precedes the first appearance of buds on mother cells in the culture (the latter coinciding with the initiation of DNA replication). In contrast, *cdc28-4* cells do not produce a significant amount of activity until 100 min at 23 °C and hardly any at all at 37 °C (see Fig. 3B). This mutant is therefore rather defective at 23 °C and more markedly so at 37 °C in producing *HO* endonuclease in these conditions. Inspection of the cellular morphology of *cdc28-4* shows that this mutant is also rather defective in cell cycle progression at 23 °C compared with the wild type; for instance, there is a much longer delay before the cells start to bud after inoculation. Figure 3C, D show similar experiments with *cdc4* and *cdc15*. In the latter, activity appears, as in the wild type, by 40 min at 37 °C, then declines to an undetectable level by 120 min. A similar pattern is seen in *cdc4*, except that although the activity declines after a peak at 60–80 min, it does not disappear afterwards.

I conclude that the appearance of *HO* endonuclease depends on execution of the *CDC28* step but not on that of *CDC4* (these are sequential steps in G_1). The endonuclease is therefore produced in early G_1 but is highly unstable as it is destroyed by the time the cells have reached mitosis.

HO transcription during outgrowth of *CDC*⁺ and *cdc*⁻ mutants

About 750 nucleotides of an 872-nucleotide single-stranded probe containing *HO* sequences¹⁰ from *Bam*HI to *Hind*III are protected from S_1 endonuclease by pre-hybridization with yeast RNA (see Fig. 4 legend). Figure 4 shows the temporal profile of *HO* RNA (as measured by the level of the S_1 -generated 750-nucleotide fragment) for G_0 cultures of various *cdc* mutants inoculated into fresh medium at 37 °C. In the wild type (Fig. 4A, right), *HO* RNA appears by 30 min, rises to a peak level by 45 min, declines to a trough at 90 min, then rises again (with the onset of a second cell cycle). The appearance of *HO* RNA throughout this period is completely blocked by addition of α factor to 40 μ g ml⁻¹ at the time of inoculation (Fig. 4A, left) or by the presence of the *ts cdc28-4* mutation (Fig. 4B). In contrast, the *cdc4-1* mutation does not block the appearance of *HO* RNA (Fig. 4C), nor does *cdc15-2* (Fig. 4D), *cdc7-1* or *cdc8-1* (results not shown). In *cdc15-2* and *cdc8-1*, *HO* RNA declines to an undetectable level by 150 min. In *cdc4-1*, on the other hand, the peak at 45 min is followed by a decline to a level significantly above background.

The similarity of the patterns obtained for *HO* RNA and endonuclease suggests that the former determines the latter.

Comparison between mother and daughter cells

Mother and daughter cells in G_0 are sufficiently different in size for it to be possible to separate them according to their sedimentation velocity. In practice, it is easier to produce a population of pure daughters than it is to produce one of pure mothers. Therefore, I have compared the ability of G_0 daughters and G_0 mixed mothers and daughters to produce either *HO* RNA or endonuclease during their traverse of the first cell cycle after inoculation into fresh medium. Daughters will obviously

become mothers after one cell cycle, but I have prevented this from happening by using a *ts cdc15* mutation. We already know that G_0 *cdc15* cells (which are a mixture of mother and daughters) will produce a single burst of *HO* RNA and endonuclease on inoculation into fresh medium at 37 °C. The question is what happens in a population of pure daughter cells? Figure 5B shows that *cdc15* daughter cells produce very little, if any, *HO* endonuclease as they traverse from G_0 to the end of mitosis. *HO* endonuclease appears, however, if daughter cells, instead of being arrested at mitosis, are allowed to enter the next cell cycle, their first as mothers (lane marked 23 °C).

Mixing of *HO* endonuclease-deficient extracts from *cdc15* daughter populations with peak extracts from *cdc15* mixed mother/daughter populations showed that endonuclease activity is proportional to the amount of the latter added to the former (results not shown). This means that the absence of activity in daughter extracts is not due to an inhibitor (or protease?) which can act in *trans in vitro*. The different behaviour of mother and daughter cells is also seen if the level of *HO* RNA is measured (Fig. 5A).

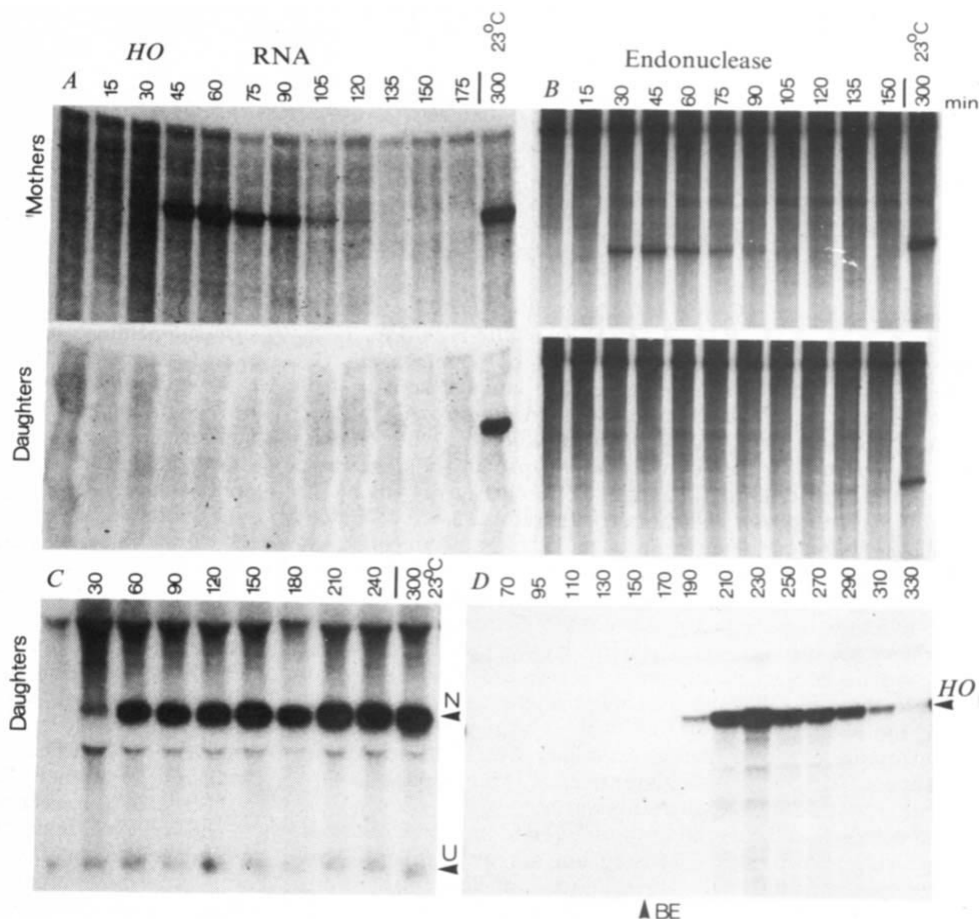
The cell cycle regulation of *HO* RNA and its confinement to mother cells is not observed for other RNA species. In Northern blots¹¹, probes prepared from the 2.5-kilobase (kb) *Hind*III fragment containing *HO* detect not only the 2-kb *HO* transcript but also a slightly larger one from a neighbouring gene (R. Jensen, personal communication). Figure 5C shows that this RNA species (marked N), in contrast to *HO*, is present throughout the cell cycle of daughter cells. The same is true of the *URA3* transcript (marked U).

Induction of *HO* RNA depends neither on bud emergence nor on spindle pole body duplication

Figure 2 shows that at least three independent pathways of cell cycle development emanate from the execution of the *CDC28* or α factor-sensitive phase of the cell cycle: (1) that leading to bud emergence (dependent on *CDC24*); (2) that leading to DNA replication (dependent on *CDC4*); (3) and that leading to spindle pole body duplication (dependent on *CDC31*). The fact that G_0 *cdc4* (but not *cdc28*) cells produce *HO* RNA on inoculation at 37 °C is consistent with an induction either before the branchpoint (Fig. 2A), on the bud emergence pathway (Fig. 2B), on the DNA replication pathway (D), or on the spindle pole body pathway (C). It would be interesting to know, therefore, whether the appearance of *HO* RNA depends on *CDC24* or *CDC31*. Experiments such as those described in Fig. 4 legend have shown that it is not dependent on *CDC24* (results not shown). A dependence on *CDC31*, however, is harder to test. Curiously, all known *ts* alleles of *cdc31*, of which there are now several, fail to arrest within one cell cycle after a shift to the restrictive temperature⁹. In fact, microscopic analysis of individual G_0 *cdc31-1* cells on germination on YEPD₍₁₎ plates at 37 °C reveals that each cell will give rise to two and only two cells having the 'dumb-bell' terminal phenotype (unlike *cdc7-1*, *cdc8-1* and *cdc15-2* which give rise to a single cell with a dumb-bell phenotype); *cdc31-1* thus undergoes only one division at 37 °C if started in G_0 . This means that we can test the dependence of *HO* RNA induction on *CDC31* by comparing the appearance of *HO* RNA in G_0 *cdc31-1* daughters inoculated into fresh medium at 37 °C with that at 23 °C. *HO*

Fig. 5 Differences between mother and daughter cells. **A**, A comparison of the level of *HO* RNA (S_1 mapping) in mixed mothers and daughters (above) and daughters (below) as they traverse from G_0 to mitosis (strain S4 at 37°C). **B**, Same as **A** but comparing the level of *HO* endonuclease; a different culture was used. **C**, Measurement of the level of the RNA transcript neighbouring *HO* (N) and the *URA3* transcript (U) during the outgrowth of daughter cells from G_0 to mitosis (Northern blotting; strain S4 at 37°C). **D**, Level of *HO* RNA during the outgrowth of G_0 *cdc31-1* (S5) daughter cells at 37°C. These cells undergo one division at 37°C before failing to duplicate their spindle pole bodies during the second. *HO* RNA is not produced during the first (daughter) division but is produced during the second (mother) division. The time by which half of the daughter cells had produced a bud (during the first division) is marked BE. Although not shown, the peak level of *HO* RNA produced in the second cycle of *cdc31-1* cells at 37°C is identical to that produced at 23°C.

Methods: G_0 cultures were prepared as described in Fig. 3 legend. Cells from 20 plates were collected by centrifugation, resuspended in a total volume of 12 ml of 0°C YEPD, and passed three times through a 21-gauge needle; 4 ml of cell suspension were then loaded on top of each of four 80-ml 10–40% sorbitol gradients (in YEPD at 10°C), and centrifuged for 7 min at 2,500 r.p.m. in an MSE 6L. Daughter cells were isolated from the top of the gradients and mixed mother/daughter fractions from throughout the gradients. Cells were then diluted threefold in 0°C YEPD, collected by centrifugation, and inoculated into 100 ml of YEPD at 37°C or 23°C. Extracts and RNA samples were prepared as described previously. The mother/daughter compositions of these cultures were analysed by washing an aliquot of cells in distilled water, staining for 5 min in 100 $\mu\text{g ml}^{-1}$ Calcafluor, and viewing with a fluorescence microscope ($\times 100$ objective) with a rhodamine filter. Such analysis suggested that no more than 1% of the cells from the daughter fraction were in fact contaminating mother cells. Mother and daughter populations were sampled at identical times. The samples marked 23°C are cells incubated at 23°C for 300 min. *HO* endonuclease was assayed as described in Fig. 3 legend. *HO* RNA was measured by S_1 mapping as described in Fig. 4 legend. The transcript neighbouring *HO* (marked N) and *URA3* (U) were measured (C) by Northern blotting¹¹. The probe used was the YIP5 plasmid (pBR322+*URA3*; ref. 25) containing the 2.5-kb *HindIII* fragment of *HO*.



RNA will not appear in the first cycle due to the cells being daughters and may or may not in the second cell cycle when the *cdc31* defect becomes apparent. Figure 5D shows that *HO* RNA does in fact appear during the second cell cycle of *cdc31-1* daughters at 37°C (note also its complete absence during the first cell cycle). This implies that appearance does not depend on spindle pole body duplication.

Regulation of the level of *HO* RNA can explain the switching pedigree

From these results I conclude that *HO* RNA and endonuclease appear only during a brief period during the G_1 phase of mother cells and not at all during the cell cycle of daughter cells (see Fig. 6A). The evidence for the first of these assertions comes from the pattern of *HO* RNA and endonuclease levels as various *cdc* mutants grow out from G_0 to their arrest point. The appearance of *HO* RNA and endonuclease is blocked by α factor (in *a* cells) or by the *cdc28-4* mutation but not by mutations in *CDC4*, 7, 8, 15, 24 or 31. These results imply that *HO* expression is cell cycle-dependent as cells blocked by α factor or by *cdc28-4* are not generally deficient in RNA transcription or protein synthesis^{12,13} and since the appearance of *HO* RNA and endonuclease is only transient in cultures of *cdc8*, 15 and 31 (that is, cells blocked at the arrest point of

cdc8, 15 and 31 do not possess *HO* RNA). It should be stressed that these results do not pinpoint the actual time of appearance of *HO* gene products during the cell cycle but rather its causal dependency on events whose timing is known. It is not possible, therefore, to state exactly when in G_1 the RNA appears or when it disappears.

The evidence for the assertion that daughter cells behave differently from mother cells, comes from experiments in which these two types of cell are separated according to their size. The induction of *HO* RNA during the cell cycle of daughter populations is less than 1/50th of that in mother/daughter populations. It is hard to exclude the possibility that the different behaviour of mother and daughter populations is due not to their being mothers as opposed to daughters so much as their cells being of different sizes. There is, however, one argument against this. Cells (mothers) do not produce *HO* RNA or endonuclease until they have executed the *CDC28* step, the stage in the cell cycle at which cells become committed to completing the mitotic cycle. This step is thought, therefore, to be regulatory and has been called START⁹. Consistent with this is the observation that cells will only execute START if they have attained a certain cell mass or size⁹. This means that mother and daughter cells will have similar sizes at START despite having very different sizes at birth. This implies that if *HO* RNA were to appear in daughter cells, it would do so when they were a similar size to mother cells which produce

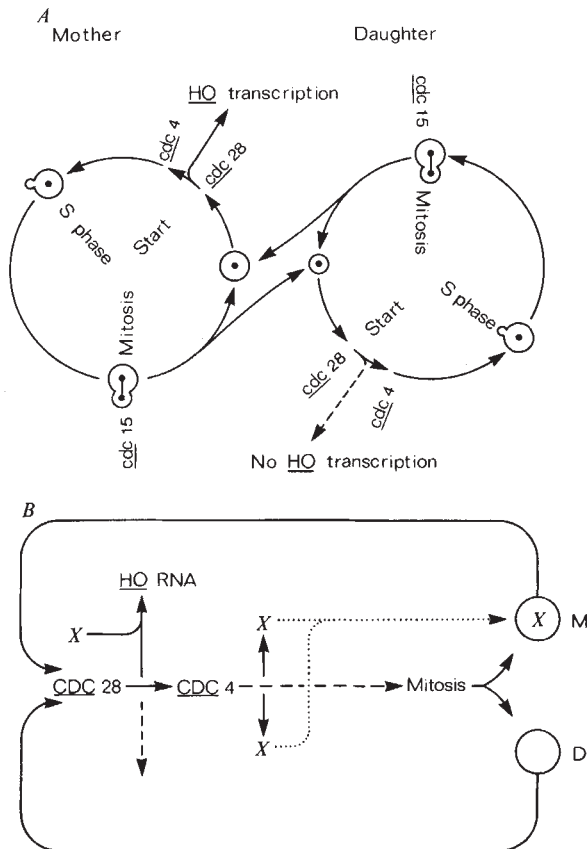


Fig. 6 *HO* expression during the cell cycle of mother and daughter cells. **A**, Differences in the cell cycle of mother and daughter cells. In the former, *HO* RNA appears during a brief phase of the cell cycle. Its appearance depends on execution of *CDC28* but not on *CDC4*. I have assumed in this diagram that transcriptional control is responsible for the level of RNA (although processing has not been ruled out). In daughter cells, *HO* RNA does not appear until they have completed the cell cycle and enter the next as a mother cell. **B**, A formal model for *HO* control. It is assumed that two factors are necessary for the appearance of *HO* RNA: (1) a cell cycle event which depends on execution of *CDC28* and (2) the presence of an unknown factor or structure called *X* which is synonymous with competence. *X* is produced in both mother and daughter cells at a stage in the cell cycle which is after (1) and is always segregated to the mother cell upon mitosis or cytokinesis. The model makes no claim to identify *X*. The *HO* gene is not only regulated by cell cycle events but also by cell type, being repressed in *a/a* diploids¹⁰. It is unknown how these two modes of regulation interact. One possibility would be for the *a/a* state to repress *X*.

HO RNA. This argument does not of course rule out the possibility of the nucleus to cytoplasmic ratio at birth being an important parameter for *HO* expression.

Whether or not *HO* is actually the structural gene for the endonuclease or only a gene whose product is a positive regulator for the latter, it is clear that the observed pattern of *HO* RNA (it is made on execution of *START* in mother cells only) can explain the observed pattern of endonuclease activity, which in turn can explain both aspects of the pattern of mating type switching. The observation that both progeny of a switching mother cell always have the new mating type can be explained easily if the *HO* endonuclease which appears on execution of *START* decays before *MAT* DNA is replicated (thus excluding the independent switching of sister loci). If this is correct, the apparent transgression of this rule by *HML* and *HMR* when they switch is *Sir*⁻ strains¹⁴ suggests that they may replicate earlier than *MAT*. The observation that daughter cells never switch is explained by the decay of all *HO* RNA and endonuclease by the time of mitosis in mothers and the inability of daughter cells to produce any on their passage through *START*.

How is the *HO* gene or its RNA regulated?

The above raises the following questions: (1) Is regulation exerted at the level of *HO* transcription? (2) If so, what is the cell cycle signal which turns on the *HO* gene (or its RNA) in mothers and why does this gene and not others require this signal to be activated? (3) Why is the gene (or its RNA) not activated during the apparently identical cell cycle of daughter cells?

An answer to the first question must probably await an analysis of the chromatin structure of *HO* or appropriate promoter fusions. Whatever the answer, the second question is central to any understanding of the molecular events which constitute commitment to or initiation of a new cell cycle. The *HO* transcript is not unique in being under cell cycle regulation. The genes for histones H2A and H2B are also activated in *G*₁¹⁵. In their case, the proximity (to the gene) of specific *ARS* sequences²⁶ has been implicated in the activation. An *ARS* sequence also flanks or is situated within the *HO* gene¹⁰ and it is therefore possible that both sets of genes are activated coordinately by a general cell cycle signal.

The cell cycle signal responsible for activating the *HO* gene (or its RNA) has been narrowed down to an event which is: (1) dependent on previous mitosis; (2) dependent on *START*; and (3) not dependent on *CDC4* (a *G*₁ event leading to DNA replication), *CDC24* (a *G*₁ event leading to bud emergence), or *CDC31* (an event leading to spindle pole body duplication). This location, though more precise, is consistent with that known for histone genes (their activation has been shown to be blocked by α factor but not by *cdc7*). In connection with the possible role of *ARS* sequences in cell cycle specific gene activation, it is noteworthy that both *HO* and histones can be activated before initiation of DNA replication. Activation may be achieved by an alteration in the structure or configuration of *ARS* sequences which is a prelude to their (possibly) acting as origins of replication. The question remains as to what this structural transition might be. One possibility is a change in the form of attachment to the nuclear membrane. This would be consistent with the observations that membrane-attached folded chromosome complexes undergo a structural transition on entering *G*₁ from *G*₀¹⁶ and that the autonomously replicating *2μ* circle is found associated with folded chromosome complexes in *G*₁ and *G*₂ but not in *G*₀¹⁷.

The third question—why the *HO* gene (or its RNA) is not activated in daughter cells—reminds one of a recurrent theme in the development of multicellular organisms, that is, how does a single cell give rise to progeny which behave differently from each other? A formal description of this process is given in Fig. 6B, where it is supposed that two factors are necessary for the activation of the *HO* gene (or its RNA): (1) passage through the *CDC28* step (*START*) and (2) the presence of an unknown factor or structure designated *X* (which is synonymous with competence⁶). *X* is produced at a stage in the cell cycle which is after *START*, and although it is produced during the cell cycle of both mother and daughter cell, it is always segregated to the mother cell on mitosis or cytokinesis. Daughter cells are therefore born without *X* and cannot produce any until after the only stage in the cell cycle at which *X* can potentially activate *HO*. The question then is what might *X* be (there is at present no good reason to suppose that *X* is a positive effector rather than the absence of a negative effector)? Does it represent the presence of a particular macromolecule which is synthesized or modified after *START* or does it represent a macromolecular structure which is assembled after *START*, in either case being segregated into the mother cell? As already discussed, *X* might reflect the nucleus to cytoplasmic ratio which is different at the birth of mothers and daughters. Alternatively, it could represent a protein which is made in *G*₁, binds to *HO* DNA, and is segregated conservatively to the chromatid which is segregated into the mother nucleus. A more concrete possibility is suggested by the observation that autonomously replicating plasmids (without centromeres) more frequently fail to be inherited by daughter cells than by mother

cells (A. Murray and J. Szostak, personal communication). One interpretation of this observation is that the parental plasmid molecules are attached in G_1 to a 'mother' compartment of the nuclear membrane and that replication produces a 'parental' chromatid which is still 'attached' and a nascent chromatid which is unattached. This interpretation raises the possibility that X might be a state of attachment of the *HO* *ARS* sequence to the nuclear membrane.

One of the many questions to be asked about X concerns its stability: for example, none of the four spores from an ascus are competent to switch in their first division, which implies that competence to switch mating type is destroyed either by the a/α state or by meiosis and sporulation (which involve starvation). It is worth stressing, therefore, that starvation *per se* (as performed in the production of all our G_0 populations)

does not destroy competence, though it may, however, reduce it (see Fig. 3A). As a mutation which destroys our hypothetical factor X should prevent *HO* transcription, it is worth mentioning that the *swi 1-1* mutation (which reduces homothallic mating type interconversion¹⁸) causes a 5–10-fold lowering of *HO* RNA levels (results not shown and M. Stern and R. Jensen, personal communication). *SWI1* might therefore encode X or a gene which affects it.

I thank Jim Hicks for providing plasmid clones containing *HO* DNA, John Yocum for supplying pB31, Rob Jensen for communicating unpublished results concerning the identity of the *HO* gene, its transcript, and others within its vicinity, and Keith McKenney for sequence information. I also thank Hugh Pelham, Jonathan Hodgkin and Cynthia Kenyon for helpful comments concerning the manuscript.

Received 20 December 1982; accepted 10 February 1983.

1. Mortimer, R. K. & Hawthorne, D. C. in *The Yeasts* Vol. 1 (eds Rose, A. M. & Harrison, J. S.) 385 (Academic, New York, 1969).
2. Hawthorne, D. C. *Abstr. Proc. 11th int. Congr. Genet.* 7, 34–35 (1963).
3. Herskowitz, I. & Oshima, Y. in *Molecular Biology of The Yeast Saccharomyces* (eds Strathern, J. N., Jones, E. W. & Broach, J.) (Cold Spring Harbor Laboratory, New York, 1981).
4. Nasmyth, K. A. *A. Rev. Genet.* 16, 439–500 (1982).
5. Hicks, J. B. & Herskowitz, I. *Genetics* 83, 245–258 (1976).
6. Strathern, J. N. & Herskowitz, I. *Cell* 17, 371–381 (1979).
7. Strathern, J. N. *et al. Cell* 37, 183–192 (1982).
8. Nasmyth, K. A. *Cell* 30, 567–578 (1982).
9. Pringle, J. R. & Hartwell, L. H. in *Molecular Biology of the Yeast Saccharomyces* (eds Strathern, J. N., Jones, E. W. & Broach, J.) 97–142 (Cold Spring Harbor Laboratory, New York, 1981).
10. Jensen, R., Sprague, G. F. Jr & Herskowitz, I. *Proc. natn. Acad. Sci. U.S.A.* (in the press).

11. Thomas, P. S. *Proc. natn. Acad. Sci. U.S.A.* 77, 5201–5205 (1980).
12. Throm, E. & Duntze, W. *J. Bact.* 104, 1388–1390 (1970).
13. Hartwell, L. H. *J. Bact.* 115, 966 (1973).
14. Klar, A. J. S., Strathern, J. N. & Hicks, J. B. *Cell* 25, 517–524 (1981).
15. Hereford, L. M., Osley, M. A., Ludwig, J. R. & McLaughlin, C. S. *Cell* 24, 367–375 (1981).
16. Piñon, R. *Chromosoma* 67, 263 (1978); 70, 337 (1979).
17. Taketo, M., Jazwinski, S. M. & Edelman, G. M. *Proc. natn. Acad. Sci. U.S.A.* 77, 3144 (1980).
18. Haber, J. E. & Garvik, B. *Genetics* 87, 33–50 (1977).
19. Crandall, M. R., Egel, R. & MacKay, V. L. *Adv. microb. Physiol.* 15, 307–98 (1977).
20. Tatchell, K., Nasmyth, K. A., Hall, B., Astell, C. R. & Smith, M. *Cell* 27, 25–35 (1981).
21. O'Farrell, P. *BRL Focus* 3(3), 1 (1981).
22. Messing, J., Crea, R. & Seeburg, P. H. *Nucleic Acids Res.* 9, 309–321 (1981).
23. Duckworth, M. L. *et al. Nucleic Acids Res.* 9, 1691–1706 (1981).
24. Nasmyth, K. A., Tatchell, K., Hall, B. D., Astell, C. & Smith, M. *Nature* 289, 244–250 (1981).
25. Scherer, S. & Davis, R. W. *Proc. natn. Acad. Sci. U.S.A.* 76, 4951–4955 (1979).
26. Osley, M. A. & Hereford, L. *Proc. natn. Acad. Sci. U.S.A.* 79, 7689–7693 (1982).

Three-dimensional architecture of a polytene nucleus

David A. Agard & John W. Sedat

Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143, USA

The three-dimensional chromosome topography in an intact nucleus has been determined using fluorescently stained Drosophila polytene chromosomes, optical fluorescence microscopy and newly developed, generally applicable, cellular image reconstruction techniques. The folding pattern is a complex mixture of parallel chromosomal segments and intertwined coils and shows extensive interaction of the chromosomes with the nuclear envelope.

ALTHOUGH much has been learned about the structure of chromatin at the levels of greatest resolution^{1–8}, surprisingly little is known of its higher-order organization. Several investigators have examined the highest levels of chromatin organization in mitotic chromosomes, and both radial-loop⁹ and sequential-coiling^{10,11} models have been suggested (for review see ref. 12). Inoué¹³, and our own work^{14,15}, suggest that in native preparations of insect sperm heads the highest level of DNA organization is a double-helical arrangement of an approximately 100-nm diameter fibre. The arrangement of chromosomes within the cell nucleus is even less well understood than the structure of the chromosomes themselves. This is primarily due to the previous lack of general methods for *in situ* analysis without disrupting the existing organization.

The notion that chromosomes are not randomly distributed within the nucleus goes back to the early 1900s, when Boveri postulated that the reproducible positioning of unique visual markers (for example, nuclear lobes in *Ascaris megacephala*) was due to a determined nuclear arrangement¹⁶. Although this idea fell into disfavour for a period, sufficient evidence has accumulated recently to help reestablish it. A nonrandom distribution of chromosomes on the metaphase plate has been reported in many species from plants¹⁷ to man¹⁸ (for review see ref. 19). Several authors have reported the specific end-to-

end attachment or association of telomeres during prophase (see, for example, refs 20, 21), as well as a spatial polarization—centromeres and telomeres at opposite poles of the nucleus. The centromeres are generally grouped together during prophase and the degree of association appears to be under direct genetic control²².

The chromosomal arrangements seen during prophase and metaphase have been postulated to be the remnants of (and consequently an argument for) chromosomal order in interphase. Indirect evidence from a variety of different experiments supports this hypothesis. Interphase chromosomes are attached to the nuclear envelope via specific chromosomal segments (centromeres, telomeres, and possibly other regions^{23–26}). There is also some recent evidence that transcriptionally active genes are closely associated with the nuclear membrane²⁷. More directly, analysis of the nonrandom distribution of radiation and mutagen induced chromosomal exchanges in somatic plant cells indicates that both homologous and nonhomologous chromosomes are nonrandomly arranged in interphase, and that this arrangement is not inconsistent with that observed during metaphase^{28,29}. All of this evidence (much of it circumstantial) points to the existence of an ordered interphase arrangement of chromosomes.

That chromosome organization is important for gene

expression is suggested by genetic studies. Perhaps the best example is the interaction of the *zeste* and *white* loci in *Drosophila*. Jack and Judd (ref. 30, see also ref. 31) reported that the *zeste* locus functions as a repressor of the *white* locus, and that repression only occurs if the two *white* alleles are physically adjacent or paired. The strength of repression was found to diminish with increasing separation. This suggests that there may be a close and defined association between loci on separate chromosomes, and that communication of regulatory signals may be spatially restricted.

In an attempt to understand the complex interactions between structure and function taking place at the chromosomal level, we have embarked on the direct determination of the spatial arrangement of *Drosophila melanogaster* chromosomes at different points in the cell cycle, and from nuclei in different developmental and transcriptional states. We report here a determination of the three-dimensional topography of chromosomes in an intact interphase nucleus. For these studies we chose polytene nuclei from *D. melanogaster* salivary glands, which are composed of cells in a well defined developmental state and so all with a structural comparison of functionally equivalent nuclei. Each of the polytene chromosomes is composed of ~1,024 haploid copies of the genome, interphase-arrested and arranged so that the genetic loci along the length of the chromosome are all in register³². This precise alignment gives rise to characteristic banding patterns that serve as markers for defined genetic loci. These greatly enlarged polytene chromosomes (3–4 µm diameter) are readily visualized by optical methods.

The topographical analysis reported here was made possible by a combination of DNA-specific fluorescent staining, optical microscopy and digital reconstruction together with enhancement techniques. This approach, essentially tomography at the cellular level, allows the structure and organization of entire cells, organelles, or their subfractions to be reconstructed in three-dimensions in essentially *in vivo* conditions. The use of optical sectioning (essentially a through-focal series) instead of physical sectioning allows a non-destructive analysis while retaining three-dimensional information. Fluorescence microscopy, using specific dyes or labelled monoclonal antibodies, allows the selective imaging of specific components in complex heterogeneous assemblies. A constrained deconvolution algorithm has been used to effectively eliminate the gaussian-emitter 'glow' effect characteristic of fluorescent imaging, resulting in a sharper, more detailed image.

Visualization of polytene chromosomes

Salivary glands from third instar *D. melanogaster* larvae were carefully removed by hand dissection and stained with the DNA-specific, non-intercalative fluorescent dye Hoechst H33258 as described elsewhere¹⁰. This dye only fluoresces when bound, and only binds to DNA. Thus only the DNA-containing structures are imaged, largely without interference from any proteinaceous or membranous structures within the cell. A non-intercalative dye was chosen to avoid chromosomal perturbations resulting from DNA-unwinding induced by intercalating dyes. Conditions were such as to minimize any artefacts arising from sequence-specific staining.

At the third instar developmental stage, the salivary glands are each composed of 140 giant cells (~70 µm diameter), arranged in a layer two cells thick. The fluorescent staining allows the polytene chromosomes within each 30–40 µm diameter nucleus to be visualized (Fig. 1). Owing to the complexity of the packing, only small segments of the individual polytene chromosomes can actually be seen in Fig. 1. Their slightly mottled appearance is the result of the barely resolved banding characteristic of polytene chromosomes. The banding pattern visualized with the Hoechst dye from fluorescent micrographs (not shown) is identical to that observed using standard squash staining procedures, which in future studies will be important for localizing specific genetic regions in the three-dimensional structure.

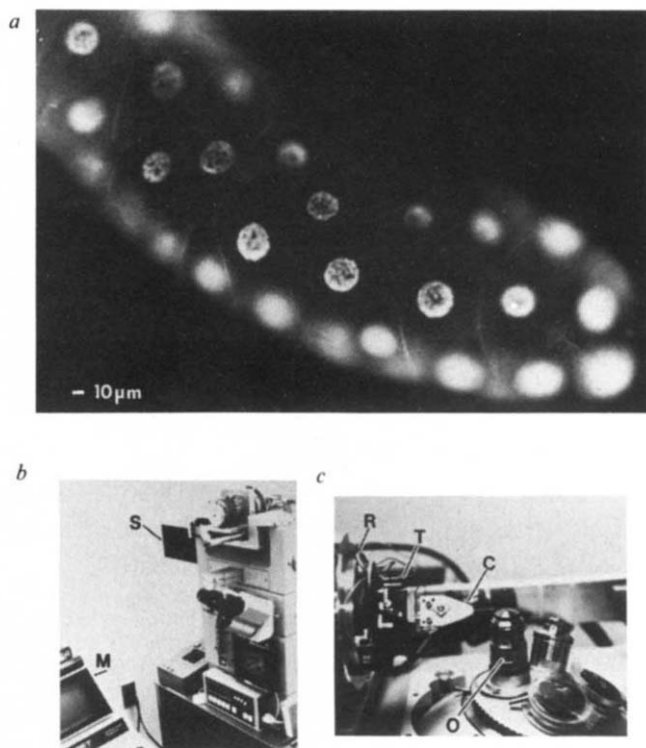


Fig. 1 *a*, An intact salivary gland from a third-instar *Drosophila* larva is shown stained for DNA with Hoechst 33258. The brightly fluorescing nuclei are clearly visible. *b*, The Zeiss Axiomat used for data collection under microprocessor control. The stepping motor connected to the focus control through an electromagnetic clutch is shown by the arrow (S), and is driven by the microprocessor (M). *c*, The tilting stage (C), has a rotation bearing (R) and translation motors (T) to align the sample mounted in a quartz capillary. The objective lens is also shown (O).

Three-dimensional data collection from thick specimens

Understanding the ways in which the individual chromosomes interact requires elucidation of the complex path followed by each chromosome through the nucleus. Mapping of this folding pattern by conventional microscopic techniques is very difficult owing to the conflicting requirements for high spatial resolution and very large depth of field. In those cases where the thickness of the object is comparable to the depth of focus, or where there is little information lost by defocus (as in electron microscopy), the object can be reconstructed from a set of tilted projections. We have used this method in the investigation of the non-crystalline three-dimensional arrangement of DNA in *Drosophila* sperm heads, which are only 0.4 µm in diameter^{13,14}.

For thicker specimens, it is necessary to reconstruct the object from a series of either physical or optical sections. Fragile structures are readily damaged by physical sectioning; and variations in section thickness, amount of shrinkage, and alignment errors can all introduce distortions into the reconstruction. By contrast, optical sectioning provides a non-destructive means of collecting three-dimensional data from thick samples and, uniquely, allows the analysis of living tissues. Nomarski optical sectioning methods have proved to be extraordinarily useful for biological studies; however, the cumulative build-up of scattering mass limits the useful specimen thickness to 15–20 µm. More seriously, such phase interference techniques provide a rather low-contrast view of the entire cellular contents. It is impossible to highlight specific macromolecular or chemically defined components.

The optical sectioning method we describe here combines the biological specificity and high contrast of fluorescence microscopy with the high resolution typical of Nomarski optics, and allows specimens with thickness considerably greater than

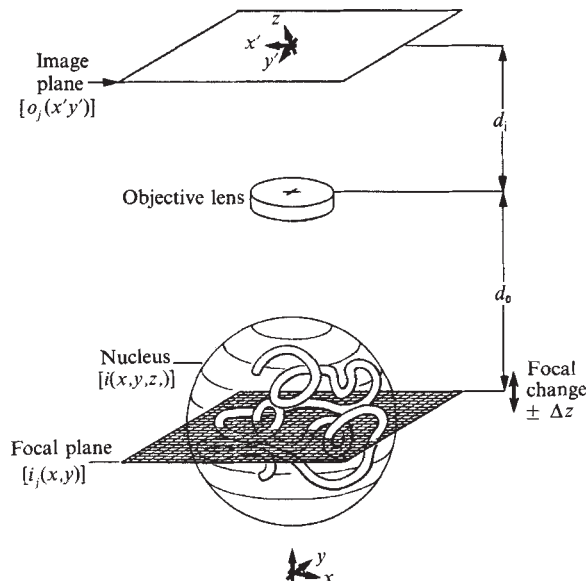


Fig. 2 A schematic representation of the optical sectioning method. The image distance d_i and the focal length of the lens f are fixed; the object distance, $d_o = fd_i/(d_i - f) = d_i/M$ (M = magnification). The image o_j is a composite of in-focus information from the focal plane i_j , and out-of-focus detail from the remainder of the specimen $i(x, y, z)$. The method described in the text allows the set of i_j (the sampled approximation of $i(x, y, z)$) to be determined from the set of observed images, the o_j .

100 μm to be readily imaged. This approach is shown schematically in Fig. 2. The thick specimen is examined using a very high resolution, narrow depth-of-focus objective lens. In these conditions only a thin plane within the sample is visualized in-focus. The recorded image represents the summation of in-focus information from that plane with out-of-focus information from the remainder of the sample, the degree of defocus varying with the distance from the in-focus plane. From a complete set of images taken at different focal points, it is possible to remove computationally from each plane the contributions of all the other planes (see below).

The accuracy of the solution depends on both the quality of the data and the accuracy with which the contrast transfer function at each level of defocus can be determined. This requires a knowledge of the microscope's optical system, as well as a precise control of the exact amount of focal shift used to record each image. This was accomplished by modifying a Zeiss Axiomat microscope so that the focus could be driven by a 1,000 step per turn stepping motor providing a change in focus of 100 nm per step. In addition, the sample (mounted in a quartz capillary on a eucentric goniometer) can be rotated to provide tilted views of the same specimen as shown in Fig. 1c. The motors, the shutter and the film advance mechanism are all controlled by microcomputer (Fig. 1b).

Three-dimensional reconstruction of a whole nucleus

To reconstruct the polytene nucleus, a series of 30 photographic images, corresponding to planes separated by 1.2 μm , was digitized on a 120- μm raster (370 nm at the sample plane) to provide a set of 30 128 $\times$ 128 pixel images. Because each negative was scanned independently, it was necessary to align each image precisely with respect to the others. The most accurate and stable procedure found was systematically to vary the translations along the X and Y coordinates, the rotation angle, and the scale factor while searching for a minimum in the normalized error function:

$$R = \frac{\sum |I - I'|}{\sum |I|}$$

where I is the reference image and I' corresponds to the scaled, rotated, translated image intensity of the image being aligned, as obtained by quadratic interpolation¹⁴. The summations are carried out only over those points where I is above some threshold (to minimize the effects of background noise on the alignment). The resultant alignment is precise to approximately 0.2 pixels, with typical R values of from 3 to 8%.

The procedure for removing the out-of-focus information from each image plane was inspired by Castleman³³, but modified so that most operations were performed in Fourier transform space, thereby greatly speeding the calculations. This approach is more exact than that used by Weinstein and Castleman³⁴ to study stained neurones; theirs is an approximate method suitable primarily for very coarse sections ($\geq 5 \mu\text{m}$ thick).

The observed image o_j for any plane j is a composite of the 'true' density distribution i_j for that plane with the blurred contributions from the m planes above and the m' planes below

$$o_j = i_j + \sum_{\substack{k=j-m' \\ k \neq j}}^{j+m} i_k * s[\Delta z \cdot (k-j)] \quad (1)$$

where $*$ represents the two-dimensional convolution operation, o_j and i_j are two-dimensional functions of the planar coordinates (x, y) , s is the defocus blurring function, and Δz is the separation between planes. The modulation transfer function s is a radially symmetric function that varies with the amount of defocus $[\Delta z \cdot (j-k)]$ and the optical parameters of the microscope. Equation (1) can be recast by taking the Fourier transform of both sides of the equation, thus converting the convolution into a multiplication

$$O_j = I_j + \sum_{\substack{k=j-m' \\ k \neq j}}^{j+m} I_k \cdot S[\Delta z \cdot (k-j)] \quad (2)$$

where O , I , and S , respectively, are the Fourier transforms of o , i , and s . Given some initial guess ($n=0$) for the I s (I^0), it is possible to generate an improved guess ($n+1=1$) (I^1) as follows

$$I_j^{n+1} = O_j - \sum_{\substack{k=j-m' \\ k \neq j}}^{j+m} I_k^n \cdot S[\Delta z \cdot (k-j)] \quad (3)$$

Thus, in an iterative fashion it is possible to develop a self-consistent solution to the three-dimensional convolution implied by equation (1). Values for the deblurred images (i_j) are determined by inverse Fourier transformation of the final I_j . The result is the determination of a set of images i_j , that when appropriately blurred and summed will equal, with arbitrary precision, the observed images o_j .

In practice the initial guess (I^0) is set equal to the observed (O). The number of sections that need be considered simultaneously ($m+m'$) depends on the choice of objective lens and the spacing between sections. With the $\times 63$ (1.25 NA) oil lens used and with $\Delta z = 1.2 \mu\text{m}$, it was only necessary to consider simultaneously a stack of seven planes (three above and three below; $m=m'=3$). Values for the optical contrast transfer function (S) for differing amounts of defocus were calculated according to the theory developed by Stokseth³⁵ using lens parameters obtained from Zeiss. In these computations the wavelength spread of the emitted light was neglected. For a higher-resolution analysis this could be included by appropriately modifying the contrast transfer function. The procedure converged to a solution after five iterations. Finally, the resultant in-focus three-dimensional image was enhanced using an iterative, constrained deconvolution technique described elsewhere^{14,36}. Such post-processing is necessary to remove the 'glow' effect characteristic of fluorescent emitters due to light scatter, and provides an apparent resolution increase of two or three-fold, without significantly increasing the noise level. The ultimate obtainable effective in-plane resolution is about 100–200 nm but was limited here to about 500 nm by the coarse scanning interval used.

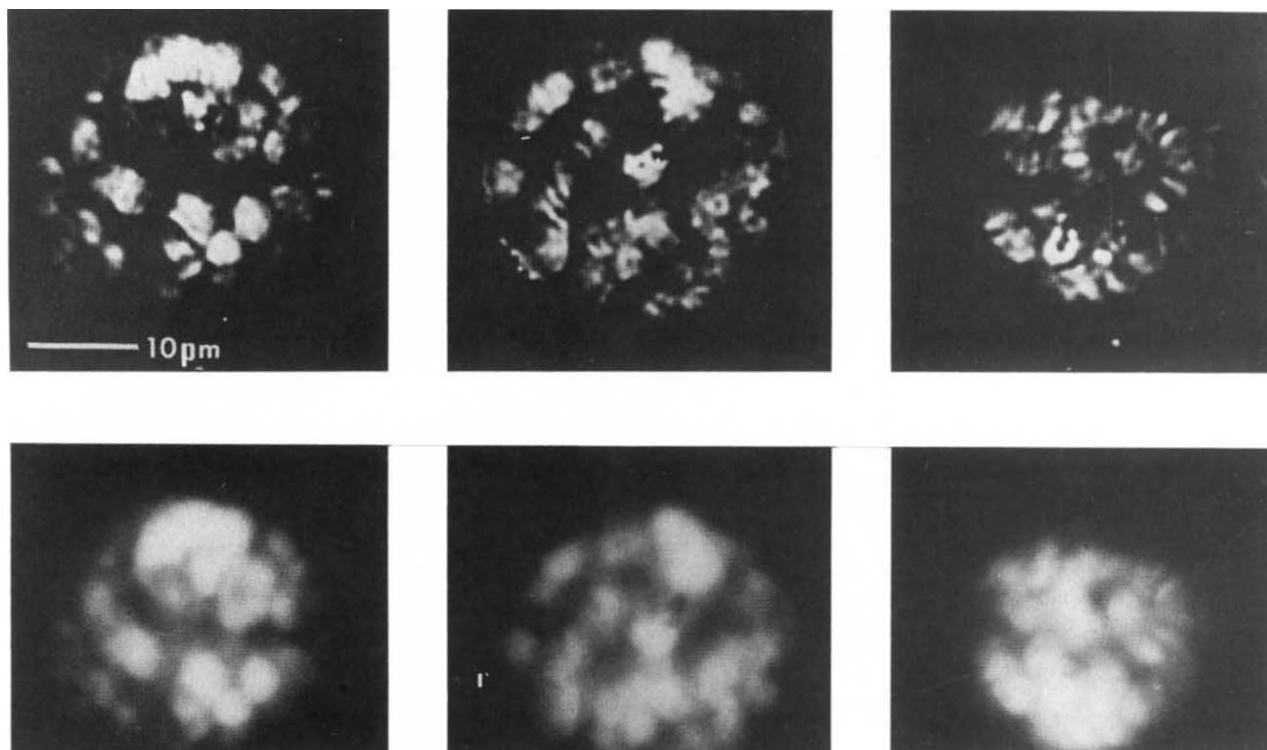


Fig. 3 Comparisons of images before (bottom) and after data processing (top) for three non-contiguous sections from the stack of 30 planes used. The data have been processed both to remove out-of-focus contamination and to remove the 'flare' due to fluorescence imaging. The major effect of the processing is to increase dramatically the definition between planes in the stack.

The result is an enhanced, three-dimensional 'DNA-density' map, which reveals the path through the nucleus followed by the DNA components within the polytene chromosomes. Before and after images of several individual sections are shown in Fig. 3 for comparison. A stereo representation of a hemisphere of the three-dimensional map viewed from the inside is shown in Fig. 4.

Chromosome topography: interpretation of the DNA-density map

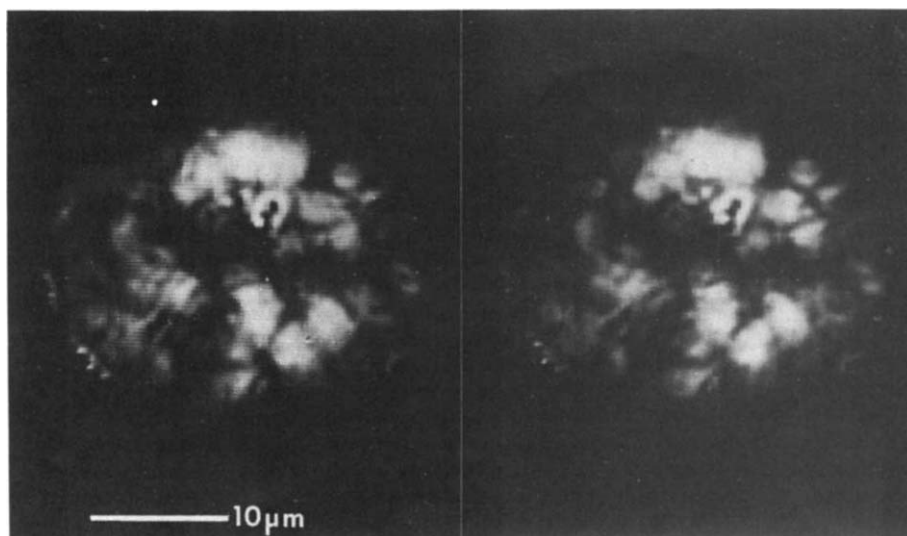
Further interpretation of the density maps shown in Fig. 4 required that the actual path followed by the polytene chromosomes be traced. This was accomplished by building a wire model using a Richard's optical comparator³⁷, as is standard in macromolecular X-ray crystallography. Computer-generated half-tone representations transferred to plastic sheets were used instead of density contour maps. Grey-level representations

were found to be significantly more interpretable than contour maps. Although only four or five sections could be overlaid at one time, the increased clarity greatly aided the model building. The physical model depicting the chromosomal folding pattern is shown in Fig. 5.

The observed folding pattern is a complex mixture of intertwined coils and parallel chromosome segments, reminiscent of folding patterns seen in proteins. These elements of chromosomal 'secondary structure' are tightly packed against the inner surface of the nuclear membrane. This close membrane association results in a predominantly chromosome-free central cavity. The nucleolus is located in this cavity, but only occupies one-quarter to one-third of the available volume.

At one end of the structure there is an extremely tight cluster of chromosomes, at what is most probably the chromocentre. Several free ends, presumably corresponding to the telomeres, can be seen at the opposite pole. These two loci define a vector

Fig. 4 A stereo perspective view of one hemisphere of the reconstructed three-dimensional 'DNA-density' map. The view is taken from the hollow central cavity (nucleolus) looking out.



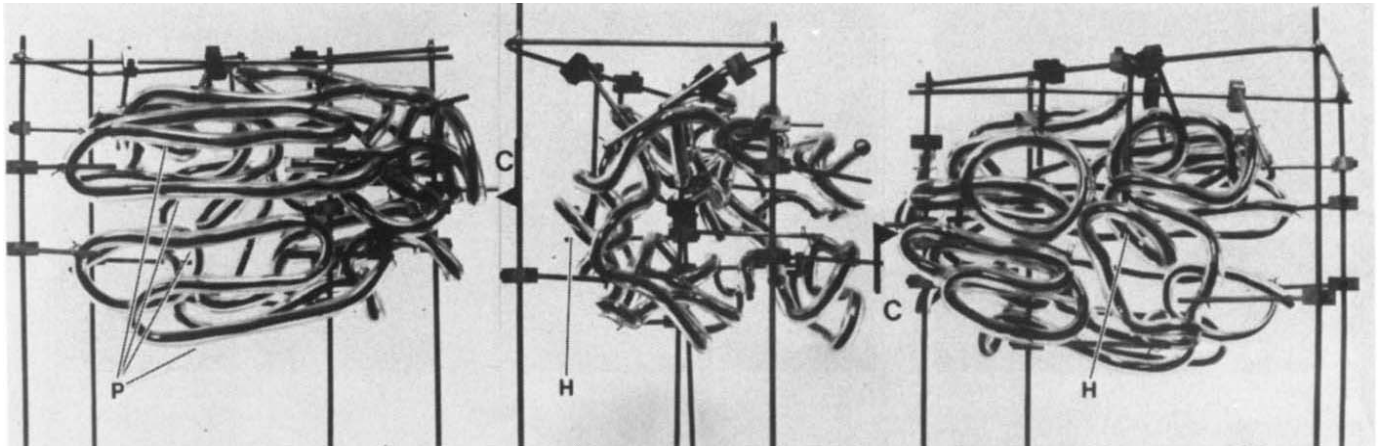


Fig. 5 Three views of the wire model constructed using a Richard's box³⁵ to superimpose the virtual image of the model onto a stack of 'DNA-density' maps mounted on plastic sheets. Note the tight cluster of chromosomes at the chromocentre (C), and the well-defined coiled (H) and parallel regions (P).

that lies perpendicular to the surface of the salivary gland, with the chromocentre positioned towards the surface. In *Drosophila* early embryonic nuclei, the individual centromeres are grouped into a densely staining structure, the chromocentre, during interphase. In *Drosophila virilis* as well as *D. melanogaster* the chromocentre has been rigorously shown to be located near the embryo surface at the cellular blastoderm stage. Barely visible fibres (the chromosomes) trail away towards the centre of the embryo, again defining a vector normal to the surface (J.W.S., manuscript in preparation). A centromere-telomere polarization has been reported in prophase plant nuclei^{20,21}. The observation that such a polarization also exists in both polytene and diploid interphase nuclei suggests that the prophase arrangement is derivative of that seen in interphase. This seemingly fundamental property (polarized chromocentre-telomeres) of interphase and prophase chromosomes may be necessary for the orderly untangling that must accompany condensation and mitosis.

Although the resolution of this reconstruction was quite adequate to determine the overall folding pattern, it was not sufficient to resolve the banding pattern in most regions of the map. As a consequence, it was not possible to identify the individual chromosome arms or to align the genetic map with the topological arrangement. In those instances where a polytene arm ran perpendicular to the image plane, the polytene chromosome was observed in cross-section, revealing that the centre of a band is relatively DNA-free, as suggested by Mortin and Sedat³⁸.

Preliminary analysis of an adjacent nucleus in the same salivary gland indicates an equivalence at least at the level of gross morphology. This was determined by using a Richard's box to optically superimpose the virtual image of the physical model on the reconstructed density map from the second nucleus. The location and relative arrangement of the major 'secondary structure' features, such as coiled or parallel regions, were found to be similar in the two nuclei, and will be reported in detail elsewhere.

Discussion

As yet, there is no physical model or conceptual framework available for integrating DNA sequence information, structural changes at the nucleosome and higher levels, and genetic data into a coherent picture of how chromosomes function within the cell nucleus. The problem of establishing such a framework is essentially a structural one. It is necessary to understand both the higher-order levels of chromatin organization and how they may be dynamically altered during transcription (as in puffing) as well as to establish the role of the spatial organization of chromosomes within the nucleus during all stages of the cell cycle. The preferential location of specific gene loci at specific

sites within the nucleus appears to have important implications for gene control and development^{30,31}.

A defined spatial arrangement can arise through specific chromosome-chromosome interactions (analogous to protein quaternary structure), through the attachment of specific chromosomal loci to specific sites on the nuclear membrane, or through both. It is quite likely that many of the observed features (such as the tight bends) result from ectopic fibre cross-linking of chromosomal segments. Cytological studies on squashed preparations have identified many specific fibre attachment points³⁹, although the number determined is probably an underestimate due to preparative damage. The demonstrated inheritance of attachment points indicates that they do not occur randomly and are under genetic control. Sedat and Manuelidis¹⁰ have shown by scanning electron microscope that the polytene chromosomes do not unfold following gentle chemical removal of the nuclear membrane, although they do collapse inwards onto the nucleolus. This suggests that the chromosomal organization is primarily maintained by chromosome-chromosome interactions, and that nuclear-membrane attachment is mainly used to support and orientate the entire assembly. The observation that the chromocentre is always orientated towards the outside of the gland or embryo indicates that the overall orientation is carefully maintained.

The determination of the limits of variability, and the correlation with factors that might influence the spatial organization requires a comprehensive analysis of a large number of nuclei. Techniques for the parameterization and refinement of the current physical model need to be developed and applied to higher resolution reconstructions. It is also necessary to align the genetic map with the observed topological folding pattern in order to correlate genetic effects at specific loci with their spatial location. This work is under way.

The study of many cellular processes such as mitosis, intracellular transport, cytoskeletal organization and rearrangements, compartmentalization and neural organization have all been hampered by the inability to analyse the relevant structures in three dimensions in living or native material. Cellular tomographic techniques (such as described here) when used in conjunction with specific dyes or monoclonal antibodies should fill this gap. The use of intensified video cameras and direct digital data acquisition should allow a three-dimensional data set to be collected either from very weakly fluorescing specimens by integration methods or very rapidly (a few seconds) from brighter samples, enabling the study of dynamic processes in live tissue.

We thank Dr R. M. Stroud for many helpful discussions and for the use of his Syntex AD1 microdensitometer and Data General Eclipse S230 computer. This work was supported by NIH grant GM25101. D.A.A. was supported by a Helen Hay Whitney postdoctoral fellowship.

Received 22 September 1982; accepted 28 January 1983.

1. Sanger, F. *et al.* *Nature* **265**, 687-695 (1977).
2. Drew, H., Takano, T., Itakura, K. & Dickerson, R. E. *Nature* **286**, 567-573 (1980).
3. Wing, R. *et al.* *Nature* **287**, 755-758 (1980).
4. McKay, D. B. & Steitz, T. A. *Nature* **290**, 744-749 (1981).
5. Anderson, W. F., Ohlendorf, D. H., Takada, Y. & Matthews, B. W. *Nature* **290**, 754-758 (1981).
6. Finch, J. T. & Klug, A. *Cold Spring Harb. Symp. quant. Biol.* **42**, 1-9 (1977).
7. Klug, A., Rhodes, D., Smith, J., Finch, J. T. & Thomas, J. O. *Nature* **287**, 509-516 (1980).
8. Kornberg, R. D. A. *Rev. Biochem.* **46**, 931-954 (1977).
9. Marsen, M. P. F. & Jaemmli, U. K. *Cell* **17**, 849-858 (1979).
10. Sedat, J. & Manuelidis, L. *Cold Spring Harb. Symp. quant. Biol.* **42**, 331-349 (1977).
11. Bak, A. L., Zeuthen, J. & Crick, F. H. C. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1595-1599 (1977).
12. Paulsen, J. R. in *Electron Microscopy of Proteins* Vol. 3 (ed. R. Harris) Ch. 4 (Academic, London, 1982).
13. Inoué, S. & Sato, H. in *Molecular Architecture in Cell Physiology* (eds Hayashi, T. & Szent-Gyorgyi, A.) 209-239 (Prentice-Hall, New Jersey, 1964).
14. Agard, D. A. & Sedat, J. W. *Proc. Soc. Photo-Opt. Instr. Engng* **264**, 110-117 (1980).
15. Agard, D. A. & Sedat, J. W. (in preparation).
16. Wilson, E. B. *The Cell in Development and Heredity* 3rd edn, pp. 215-225 (Macmillan, New York, 1925).
17. Finch, R. A., Smith, J. B. & Bennett, M. D. *J. Cell Sci.* **42**, 391-403 (1981).
18. Miller, O. J., Mukherjee, B. B., Breg, W. R. & Gamble, A. van N. *Cytogenetics* **2**, 1-14 (1963).
19. Lacadena, J. R. & Ferrer, E. *Chromosoma* **67**, 77-86 (1974).
20. Ashley, T. *J. Cell Sci.* **38**, 357-367 (1979).
21. Rabi, C. *Morph. Jb.* **10**, 214-330 (1885).
22. Thomas, H. *Chromosoma* **42**, 87-94 (1973).
23. Avivi, L., Feldman, M. & Bushuk, W. *Genetics* **62**, 745-752 (1969).
24. Sved, J. A. *Genetics* **53**, 747-756 (1966).
25. Davies, H. G., Murray, A. B. & Walmsley, M. E. *J. Cell Sci.* **16**, 261-299 (1974).
26. Murray, A. B. & Davies, H. G. *J. Cell Sci.* **35**, 59-66 (1978).
27. Robinson, S. I., Nelkin, B. D. & Vogelstein, B. *Cell* **28**, 99-106 (1982).
28. Werry, P. A. Th. J., Stoffejen, K., Engles, F. M., van der Jaan, F. & Spanjers, A. W. *Chromosoma* **62**, 93-101 (1977).
29. Evans, H. J. & Bigger, T. R. C. *Genetics* **46**, 277-289 (1961).
30. Jack, J. W. & Judd, B. H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1368-1372 (1979).
31. Bingham, P. M. *Genetics* **95**, 341-353 (1980).
32. Beerman, W. *Protoplasmatologia* **60**, 1-152 (1962).
33. Castleman, K. R. *Digital Image Processing*, 351-360 (Prentice-Hall, New Jersey, 1979).
34. Weinstein, M. & Castleman, K. R. *Proc. Soc. Photo-Opt. Instr. Engng* **26**, 131-138 (1971).
35. Stokseth, P. A., *J. opt. Soc. Am.* **59**, 1314-1321 (1969).
36. Agard, D. A., Steinberg, R. A. & Stroud, R. M. *Analyt. Biochem.* **111**, 257-268 (1981).
37. Richards, F. M. *J. molec. Biol.* **37**, 225-230 (1969).
38. Mortin, L. & Sedat, J. W. *J. Cell Sci.* **57**, 73-113 (1982).
39. Kauffman, B. P. & Iddles, M. K. *Port. Acta. Biol.* **A7**, 225-248 (1963).

The yeast tRNA^{Tyr} gene intron is essential for correct modification of its tRNA product

Peter F. Johnson* & John Abelson†‡

Departments of Biology* and Chemistry,† University of California at San Diego, La Jolla, California 92093, USA, and The Agouron Institute, 505 Coast Boulevard South, La Jolla, California 92037, USA

Precise deletion of the intervening sequence of a yeast tRNA^{Tyr} ochre suppressor gene (SUP6) significantly reduced its suppressor activity relative to that of the unaltered gene. This is probably the result of the absence of the pseudouridine modification, normally present at the middle anticodon position of mature suppressor tRNA, in tRNA synthesized in vivo from the deleted gene.

SINCE the discovery that many eukaryotic genes are discontinuous (see refs 1, 2 for reviews), a number of studies have been directed towards identifying a function for intervening sequences (IVSs). The deletion of introns from SV40 late genes^{3,4} and SV40/rabbit β -globin hybrid genes^{5,6} was shown to eliminate stable cytoplasmic mRNA production, leading to the suggestion that the processing and transport to the cytoplasm of mRNA might be obligatorily linked⁶. However, this model does not explain the synthesis of mRNA from genes that, like those for histones, lack IVSs⁷⁻⁹, and there are now examples of genes that are functional after the removal of their IVSs^{10,11}, and of stable mRNAs still containing an intervening sequence¹². A definitive role for the introns of nuclear genes in mRNA biosynthesis has therefore remained elusive, although in yeast mitochondrial mRNA genes IVSs have been found that apparently encode proteins involved in precursor mRNA processing (for review see ref. 13).

IVSs are also found in genes coding for ribosomal and transfer RNAs, although, again, little is known of their functions. We estimate that one-fifth of yeast tRNA genes contain introns, so splicing cannot be a general requirement of tRNA biosynthesis. Sequence analysis of several split tRNA genes from yeast¹⁴⁻¹⁸ has revealed the following characteristics of their IVSs: (1) they reside one base to the 3' side of the anticodon; (2) within a tRNA gene family their sequences are identical, or nearly so, but there are no homologies between families; and (3) they always contain a sequence complementary to the anticodon. Pairing between these sequences may be involved in the precursor secondary structure^{19,20} (Fig. 1). Co-purification of the yeast splicing endonuclease activities for a number of pre-tRNAs suggested that there may be a common splicing apparatus for

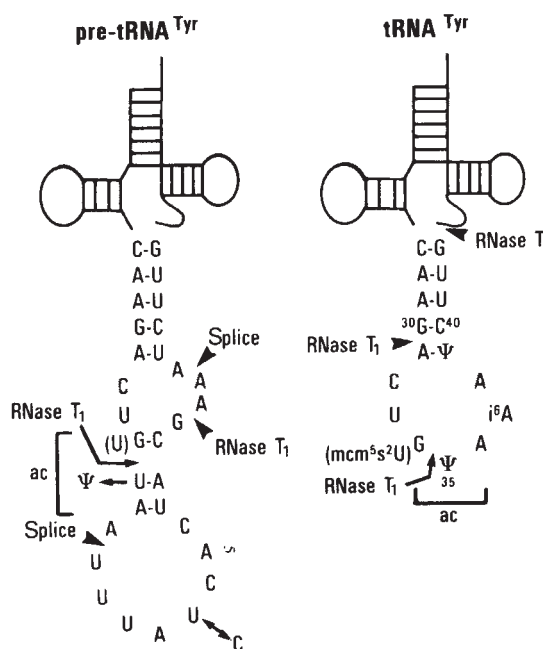


Fig. 1 Anticodon regions of pre-tRNA^{Tyr} and mature tRNA^{Tyr}. Base modifications, relevant RNase T₁ cleavage sites and splice sites are indicated. The nucleotide corresponding to the SUP6 mutation is shown in parentheses. Abbreviations: ψ , pseudouridine; i⁶A, N⁶-isopentenyladenosine; mcm⁵s²U, 5-(methoxycarbonylmethyl)-2-thiouridine; ac, anticodon. Numbering of nucleotides in the mature tRNA is according to previous convention³⁸.

‡ Present address: Division of Biology, California Institute of Technology, Pasadena, California 91125, USA.

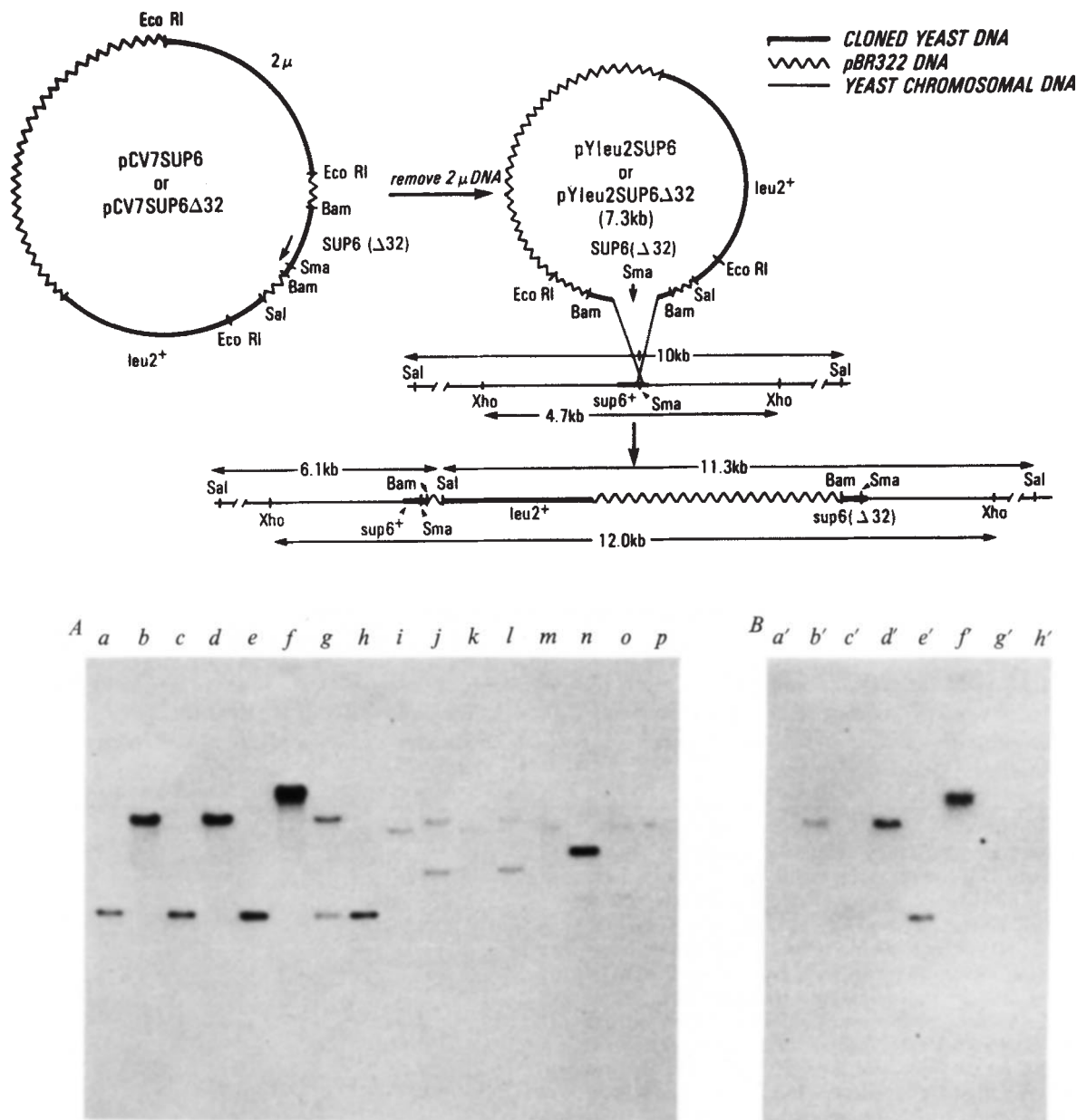


Fig. 2 Upper panel: Integration of *SUP6*-containing plasmids at *sup6*, and subsequent replacement of the *sup6⁺* allele. Gene replacement strains are indicated by the name of their parental transformant followed by -r (for example, 302f'-r26). Lower panel: **A**, Autoradiogram of a blot containing *XhoI*-digested (lanes a-h) and *SalI*-digested (lanes i-p) DNAs hybridized with a probe from the 5' flanking region of *sup6*. Yeast strains from which the DNAs were extracted: lanes a, i, DH484 (untransformed host); lanes b, j, 302f'; lanes c, k, 302f'-r26; lanes d, l, 303a; lanes e, m, 303a-r8; lanes f, n, 304b; lanes g, o, 5a; lanes h, p, 5a-r1. **B**, Autoradiogram of the same filter (*XhoI* digests only) after rehybridization with an oligonucleotide probe specific for *SUP6*Δ32.

Methods: 2 μ DNA was removed from pCV7SUP6 and pCV7SUP6Δ32²⁴ by partial digestion with *EcoRI*. The desired 7.3 kb bands were purified by agarose gel electrophoresis, circularized with T4 DNA ligase and used to transform *E. coli* (HB101), selecting for ampicillin resistance. The resultant plasmids (pYleu2SUP6 and pYleu2SUP6Δ32) were purified, and 10 μg of each was digested with *SmaI*, followed by phenol extraction and ethanol precipitation. The linear DNA was used to transform the yeast strain DH484 (a, *ade2-1*, *leu2-3*, *leu2-112*, *can1-100*, *trp5-48*, *ura4-11*, *lys1-1*) according to the method of Hsiao and Carbon³⁹. *Leu⁺* colonies were selected (5–30 transformants per μg DNA) and tested for suppression of the *ade*, *trp* and *lys* ochre mutations. Approximately 20% of the *Leu⁺* transformants also exhibited suppressor activity. DNA was prepared⁴⁰ from some of these strains, digested with either *XhoI* or *SalI* and separated by electrophoresis on a 1% agarose gel. The DNA was then blotted to nitrocellulose²⁶ and hybridized⁴¹ with a nick-translated restriction fragment derived from the 5' flanking region of *sup6*. Strains containing a precise replacement of the *sup6⁺* gene by *SUP6* or *SUP6*Δ32 were constructed by screening for *Leu⁺* revertants in which the integrated plasmid was excised by recombination between the *sup6* sequences. Recombination was stimulated by brief exposure to UV irradiation (90% viability; R. Rothstein, personal communication) after appropriate dilution and plating of the yeast cultures, and *Leu⁺* colonies identified by replica-plating. The revertants were then tested for retention of the suppressor phenotype. In **B** label was removed from the filter by denaturation⁴⁶ and the blot rehybridized with a ³²P-labelled 21 base oligonucleotide used to construct *SUP6*Δ32²³, using conditions in which only hybridization to the deletion mutant could occur²³. Faint bands corresponding to the parental *SUP6* fragments are due to incomplete removal of the previous probe.

all pre-tRNAs²¹. Thus, the splicing enzymes probably recognize certain structural features of the pre-tRNAs rather than the RNA sequence *per se*. In light of this, the almost complete lack of sequence divergence among the IVSs of a given tRNA gene family is surprising (although it is possible that intergenic conversion mechanisms have a role in maintaining sequence homogeneity; see, for example, ref. 22). As such sequence conservation could indicate a positive role for IVSs in tRNA synthesis, we were interested to determine whether the presence of an IVS (or the act of splicing itself) is necessary for the optimal expression of a spliced tRNA species.

We previously reported the construction of a mutant tRNA^{Trp} ochre suppressor gene (*SUP6Δ32*) from which the 14 base pair (bp) IVS had been precisely deleted²³. *SUP6Δ32*, and also the parent suppressor gene *SUP6*, were cloned into the transformation vector CV7. After transformation of an appropriate yeast host, the ability of these two genes to suppress three ochre nonsense mutations was compared. No difference in phenotype could be detected. However, since these genes were present in the cell on high copy number plasmids, the possibility existed that a reduction of suppressor efficiency in the mutant was compensated by the effect of high gene dosage. To make a more sensitive comparison of these two genes, we have exploited the capability in yeast of replacing a gene in the chromosome with a cloned mutant allele²⁴. In this technique, a plasmid carrying the altered gene is first integrated at the locus to be replaced by homologous recombination, creating a structure consisting of plasmid sequences flanked by the chromosomal gene and the cloned copy (see Fig. 2). A subsequent recombination event between the two genes results in excision of the plasmid and, depending on the point of crossover, leaves either the wildtype sequence or the mutation in the genome. Using this approach to replace the resident *sup6* gene with either *SUP6* or *SUP6Δ32*, we show here that removal of the IVS converts the *SUP6* gene to a less efficient suppressor. We also demonstrate that the presence of the IVS is necessary for the modification of uridine in the tRNA anticodon to pseudouridine (ψ). We believe that the failure to make this modification is responsible for the decrease in suppressor efficiency of the mutant.

Construction of yeast strains carrying single copies of *SUP6* and *SUP6Δ32*

To create plasmids capable of integration into the yeast genome, the vector-derived 2 μ DNA was removed from the recombinant plasmids pCV7*SUP6* and pCV7*SUP6Δ32* by partial digestion with *EcoRI*. The resultant plasmids (pYleu2*SUP6* and pYleu2*SUP6Δ32*) contain pBR322, the yeast *LEU2⁺* gene, and the indicated suppressor gene (Fig. 2). These plasmids can integrate into yeast chromosomal DNA by homologous recombination at either the *sup6* or *LEU2* loci. As replacement of the resident *sup6⁺* gene by a cloned copy first requires integration of the plasmid at this locus²⁴, it was desirable to direct integration specifically to this site. Orr-Weaver *et al.*²⁵ have shown that the free ends of cloned yeast sequences created by restriction enzyme cleavage are highly recombinogenic. Thus, linearization of a plasmid containing multiple yeast

sequences serves to direct integration to the site of cleavage, and also increases the transformation frequency significantly. Accordingly, a unique *SmaI* site present at the 3' end of the *SUP6* gene was used to linearize pYleu2*SUP6* and pYleu2*SUP6Δ32*. The linear DNAs were used to transform the host strain DH484 (a haploid strain that carries a double mutation in the *LEU2* gene (*leu2-3*, *leu2-112*) and contains the ochre mutations *ade2-1*, *trp5-48* and *lys1-1*) and *Leu⁺* colonies were selected.

Leu⁺ transformants were tested for suppression of the *ade*, *trp* and *lys* mutations. Typically, only about 20% of the *Leu⁺* cells also carried the suppressor phenotype. That many of the *Leu⁺* cells, which contain the plasmid integrated at *sup6*, are also *Sup⁻* (non-suppressing) is presumably due to a high frequency of gene conversion associated with the integration of linear DNA molecules²⁵. Of the cells exhibiting suppression, however, there was a clear distinction between the pYleu2*SUP6* and the pYleu2*SUP6Δ32* transformants. Cells containing the *SUP6* gene were able to grow well on plates lacking adenine, tryptophan or lysine. In addition, the colonies were white, indicating efficient suppression of the *ade2-1* mutation (*ade2⁻* cells accumulate a characteristic red pigment). In contrast, the *SUP6Δ32* mutant was a much weaker suppressor, allowing normal growth in the absence of tryptophan but only weak growth without lysine and no growth without adenine. These colonies were pink, indicative of partial suppression of *ade2-1*.

Certain of the *SUP6Δ32* transformants were able to grow without adenine. One of these (304b) was picked and was shown by Southern blotting analysis²⁶ (see below) to contain a multiple tandem plasmid integration at the *sup6* locus. We estimate that this strain contains three or four copies of the plasmid (although because of possible gene conversion events one cannot be certain that this number of *SUP6Δ32* alleles are present). 304b has similar growth properties to a diploid DH484 strain bearing a single copy of the normal *SUP6* allele (designated 5a). 5a was discovered in a transformation experiment in which a spontaneously arising *a/a* diploid of DH484 was fortuitously used as the recipient strain. 5a was identified as a diploid on the basis of restriction analysis of its genomic DNA (see below). Both 304b and 5a grow poorly without adenine, and are pink and therefore do not completely suppress the *ade2-1* mutation. By equating the phenotypes and comparing the suppressor gene dosages of these two strains, one can make a rough estimate that *SUP6Δ32* has one-sixth the suppressor activity of *SUP6*.

The physical structures of several transformants were determined by restriction digestion and Southern blotting of their genomic DNAs, as shown in Fig. 2. The DNA was digested with either *XhoI*, which does not cleave within pYleu2*SUP6*, or *SalI*, which cleaves once within the plasmid. The blots were hybridized with a radioactively labelled restriction fragment containing sequences from the 5' flanking region of the *sup6* gene. Integration of the plasmid at *sup6* results in an increase in size of the hybridizing *XhoI* fragment from 4.7 kb to 12.0 kb. Thus, it can be seen that strain 302f (lane b), transformed with pYleu2*SUP6*, and strain 303a (lane d), transformed with pYleu2*SUP6Δ32*, each contain a single plasmid copy integrated at *sup6*. In lane f (strain 304b), an *XhoI* band of about 34 kb

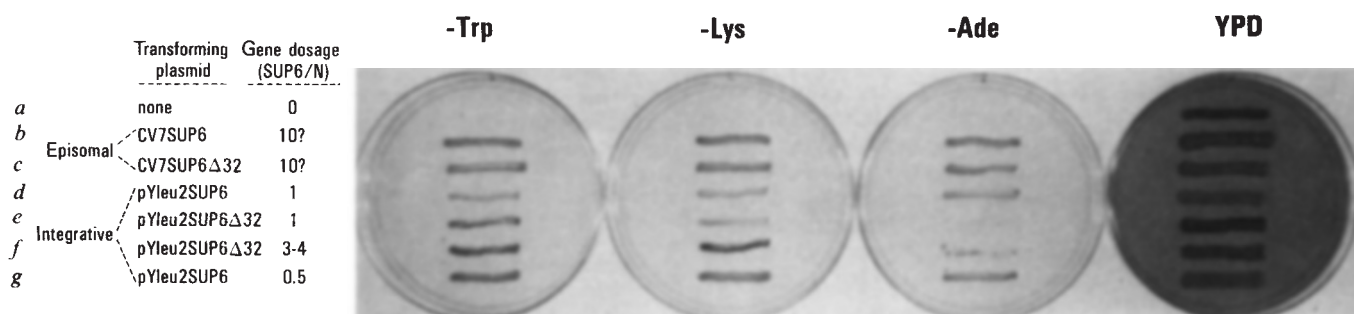


Fig. 3 Phenotypes of the yeast transformants. a, DH484; b, 3011a; c, 3012a; d, 302f-r26; e, 303a-r8; f, 304b; g, 5a-r1. Cells were streaked on defined minimal media lacking the indicated nutrient and on rich medium (YPD)⁴².

appears, consistent with the tandem integration of four plasmid copies. This was confirmed by digestion with *Sal*I (lane *n*), in which two bands corresponding to those seen for a single copy transformant are present. (The plasmids used to construct 304b and 5a contain *sup6* inserts in opposite orientation to those present in 302f' and 303a; thus these *Sal*I fragments, which represent the junctions between plasmid and chromosomal sequences, are different from those in 302f' and 303a). There is also an intense band equal in mobility to the linear transforming plasmid, which would be expected for a multiple tandem repeat. DNA from transformant 5a is shown in lanes *g* and *o*. Here there are bands representative of both the host strain and a single copy integrant. As mentioned above, we believe that this cell is a diploid with pYleu2SUP6 integrated into one of the two *sup6* loci.

The most rigorous comparison of the two alleles required the precise replacement of the resident *sup6*⁺ gene with *SUP6* or *SUP6Δ32*. Thus, any effects of flanking plasmid sequences on the expression of these genes could be excluded. This can be accomplished by a recombinational event between the two *sup6* copies in which the integrated plasmid is excised, removing the *sup6*⁺ gene and leaving behind the cloned suppressor copy²⁴. Such strains were identified as Leu⁻ revertants that retained the suppressor phenotype. The suppressor was retained in approximately 50% of the Leu⁻ colonies. These 'pop-out' strains (302f'-r26, 303a-r8 and 5a-r1) were not detectably different from their Leu⁺ parents (302f', 303a and 5a) in suppressing the ochre mutations. The growth on suppression tester plates of these and several of the other transformants mentioned above is shown in Fig. 3. The gene replacement strains were shown to have precisely excised the integrated plasmids, as the *sup6* *Xho*I fragments were restored to their original 4.7 kb size (Fig. 2A, lanes *c*, *e*, *h*).

The *Xho*I blot was rehybridized with a ³²P-labelled 21 nucleotide synthetic oligonucleotide which was used to create the IVS deletion²³ (Fig. 2B). Hybridization conditions²³ were used in which a stable hybrid could form with *SUP6Δ23* but not with *SUP6* where the duplex is interrupted by the unpaired intron. The strong hybridization in lanes *d*', *e*' and *f*' confirms that transformants 303a, 303a-r8 and 304b have obtained the *SUP6Δ32* allele.

Analysis of the suppressor tRNA

The observation that there is reduced suppressor activity from the *SUP6* mutant lacking the IVS raises questions regarding the defect in the suppressor tRNA. The absence of an intron might reduce the amount of tRNA transcription from this gene if the intron is necessary for optimal promoter efficiency. Alternatively, splicing could provide the most efficient pathway for some aspect of tRNA biogenesis such as nucleus to cytoplasm transport. Another possibility is that the amount of RNA is not affected, but some intrinsic property of the mature tRNA is dependent on the presence of an IVS at some stage during the maturation pathway. To address these questions we purified and analysed suppressor tRNA^{Tyr} from cells containing either the *SUP6* or *SUP6Δ32* genes.

Cells were labelled with ³²P-phosphate, and the low molecular weight RNA extracted. Mature tRNA^{Tyr} was purified by a modified two-dimensional electrophoresis procedure (Fig. 4A). Since there are eight tRNA^{Tyr} genes in yeast²⁷, only a minor fraction of the RNA consists of suppressor RNA. However, the amount of suppressor RNA in this mixture can be determined by taking advantage of the fact that the G to U mutation in the suppressor anticodon creates a new RNase T₁ digestion product. The 13 base oligonucleotide produced is the largest fragment generated by RNase T₁ cleavage, and can easily be isolated by electrophoresis on a 20% polyacrylamide gel (Fig. 4B, band 1). This band is not seen in RNA from the untransformed host (data not shown).

The relative amount of suppressor RNA produced in each strain was quantitated by measuring the radioactivity present in the suppressor-specific oligonucleotide. This number was

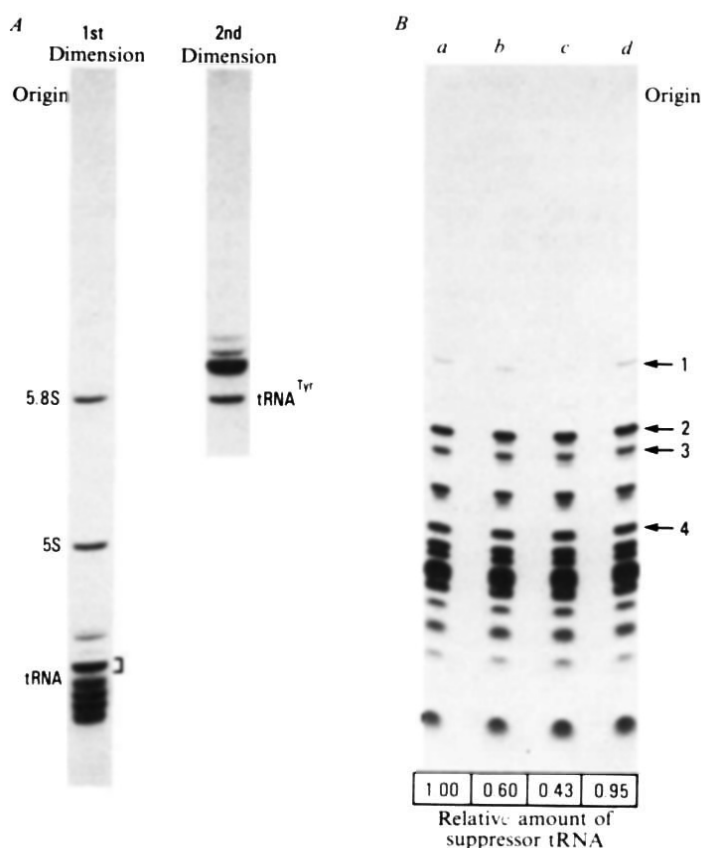


Fig. 4 **A**, Two-dimensional gel purification of ³²P-labelled tRNA^{Tyr}. The band indicated by the brackets was electrophoresed in the second dimension. The fastest migrating band in the second dimension contained tRNA^{Tyr}. **B**, Quantitation of suppressor RNA by RNase T₁ digestion. The amount of suppressor tRNA relative to total tRNA^{Tyr} was determined by measuring the amount of label in band 1 and dividing by the sum of the amounts in bands 2–4. The relative amount of suppressor tRNA is expressed as the fraction of that measured in strain 302f'-r26 (one copy of *SUP6* per haploid genome). Suppressor tRNA represents 11.4% of total tRNA^{Tyr} in 302f'-r26. Strains used: lane *a*, 302f'-r26; lane *b*, 303a-r8; lane *c*, 5a-r1; lane *d*, 304b.

Methods: 40 ml of low phosphate YEPD medium⁴³ containing 5 mCi ³²P-phosphate was inoculated with 0.8 ml of a saturated yeast culture and shaken at 30 °C for 6 h. The cells were collected by centrifugation and the low molecular weight RNA extracted and purified by DEAE-cellulose chromatography according to the procedure of Knapp *et al.*⁴⁴. The ethanol-precipitated RNA was dissolved in dye mix (95% formamide, 10 mM EDTA, 0.1% each xylene cyanol FF and bromophenol blue) and applied to a 40 cm long, 2 mm thick 15% polyacrylamide gel containing 7 M urea (20:1, acrylamide: bisacrylamide; 88 mM Tris-borate pH 8.3; 2.5 mM EDTA). Electrophoresis was conducted at 400–600 V until the xylene cyanol had migrated 1.5 times through the gel. The gel was autoradiographed and the region indicated by the bracket was excised from the gel. The gel slice was polymerized into a 20% polyacrylamide, 4 M urea gel in the orientation such that electrophoresis would continue in the same direction. Electrophoresis in the second dimension was for 48 h at 600 V. The fastest migrating band contained tRNA^{Tyr}, slightly contaminated with a second RNA species. tRNA^{Tyr} was eluted from the preparative gel⁴⁴, precipitated with ethanol and digested with RNase T₁ (5 units RNase T₁ in 10 μl of 10 mM Tris pH 7.4/1 mM EDTA with 10 μg *E. coli* tRNA carrier; 1 h at 30 °C). The digested RNA was lyophilized to dryness, dissolved in dye mix and electrophoresed on a 20% polyacrylamide sequencing gel until the bromophenol blue marker had migrated 25 cm. Band 1 is the suppressor-specific 13 base anticodon oligonucleotide and band 2 is the 9 base fragment from wild-type tRNA^{Tyr} (see Fig. 1); the identities of the other bands have not been determined.

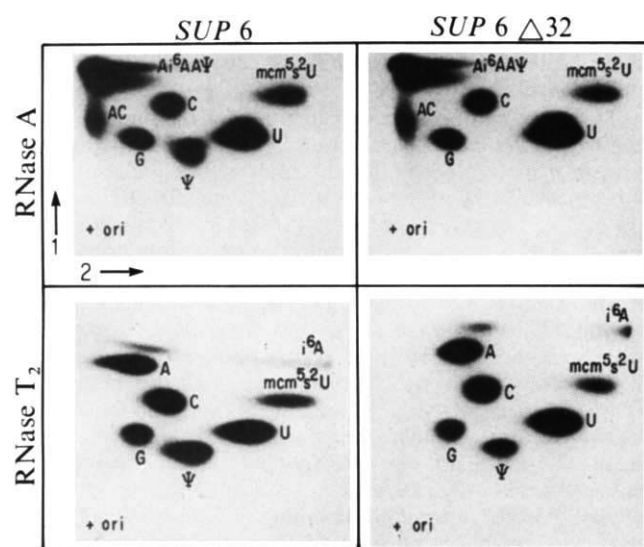


Fig. 5 Modified base analysis of anticodon oligonucleotides derived from *SUP6* and *SUP6Δ32* tRNA^{Tyr}. Band 1 RNA (Fig. 4B) from strains 302f'-r26 (*SUP6*) and 303a-r8 (*SUP6Δ32*) was eluted from the gel and ethanol precipitated with the addition of 20 μg *E. coli* tRNA carrier. The samples were each divided into two aliquots, one of which was digested with RNase A (0.33 mg ml⁻¹ RNase A in 5 μl of 10 mM Tris pH 7.4/1 mM EDTA; 4 h at 37 °C) and the other with RNase T₂ (1 unit RNase T₂ in 5 μl of 50 mM ammonium acetate pH 5.0; 4 h at 37 °C). The products were separated by two-dimensional TLC (10 cm × 10 cm cellulose plates; the solvent system of Saneyoshi *et al.*²⁸ was used). Streaking of the i⁶A spot in the RNase T₂ digests is due to breakdown of this species to A during chromatography in the second dimension.

normalized to the amount of tRNA^{Tyr} in the preparation by dividing it by the sum of radioactivity in three RNase T₁ oligomers common to wildtype and suppressor tRNA^{Tyr} (Fig. 4B, bands 2–4). The data are presented relative to strain 302f'-r26 (one copy of the *SUP6* gene) in which the normalized level of suppressor is taken to be one. Thus it can be seen that the *SUP6Δ32* gene (strain 303a-r8) produces mature tRNA at a level of about 60% that of the unaltered *SUP6* gene. While the data in Fig. 4B represent a single labelling experiment, this ratio of RNA levels is reproducible to within ±10%. RNase T₁ digests of the RNA from 5a-r1 (lane c) and 304b (lane d) are also shown. Strain 304b produces nearly the same amount of suppressor RNA as 302f'-r26, and twice the amount produced by 5a-r1, yet as described above it is a poorer suppressor than 302f'-r26 and has equivalent suppressor activity to 5a-r1. These facts strongly suggest that it is not the amount of suppressor RNA that is responsible for the difference in phenotype between *SUP6* and *SUP6Δ32*. Rather, some property of the tRNA itself is implicated as the defect caused by the absence of an IVS.

One likely candidate for a structural change affecting tRNA function would be the absence of or difference in a base modification. Because of their proximity to the IVS, modifications within the anticodon stem and loop region of tRNA^{Tyr} are of particular interest. Figure 1 shows the sequence of this area in the mature tRNA and precursor, identifying the modified bases. To analyse this region, the 13 base RNase T₁ oligonucleotides derived from *SUP6* or *SUP6Δ32* tRNA^{Tyr} were eluted from the gel and further digested with either RNase T₂ or RNase A. RNase T₂ cleaves at every base, whereas RNase A cuts after pyrimidine residues; both enzymes generate 3'

nucleotides. The digestion products were then separated by two-dimensional TLC to resolve the modified nucleotides²⁸. Autoradiograms of the TLC plates can be seen in Fig. 5. A striking difference between *SUP6* and *SUP6Δ32* RNA can be seen in the RNase A digests, where a pseudouridine mononucleotide is present in *SUP6* but is undetectable in *SUP6Δ32*. This can be explained by the lack of the normal U → ψ modification at the second position in the anticodon (position 35) in *SUP6Δ32* tRNA. The other ψ residue within the 13-mer (at position 39) remains part of the tetranucleotide Ai⁶AAψ in an RNase A digest and migrates in a different position on the plate. The RNase T₂ digest proves that the ψ at position 39 is present in *SUP6Δ32* RNA, as a spot corresponding to ψp now appears. The molar yields of the various nucleotides from these digests were determined (Table 1). It can be seen that the decrease in ψ in *SUP6Δ32* RNA is associated with a corresponding increase in the amount of U, indicating that an unmodified U is present at position 35. Furthermore, by comparing the two RNase T₂ digests, there appear to be no other differences in modified base composition. Specifically, the isopentenyl A (i⁶A) at position 37 and the U to mcm⁵s²U conversion found at the anticodon wobble position in tRNA^{Tyr} ochre suppressors²⁹ are not affected by the absence of the IVS. The failure to convert U to ψ in the anticodon of *SUP6Δ32* tRNA was observed in all yeast transformants tested, regardless of whether the gene was genomically integrated or residing on an episomal vector.

Analysis of pseudouridine in the anticodon of pre-tRNA^{Tyr}

The requirement for an IVS for production of the anticodon ψ could be explained if this modification takes place only on the intron-containing tRNA precursor. In a previous study of modified nucleotides present in pre-tRNA^{Tyr} isolated from the *ts136* mutant, the yield of ψ at the anticodon was <10%³⁰. However, the extent of base modification in precursors generated in this way is variable, and is probably sensitive to the precise labelling regimen used. We have therefore repeated the analysis of pre-tRNA^{Tyr}. The labelled RNA was digested with RNase T₁, generating a 15 base oligonucleotide that contains the potential ψ residue at its 5' terminus and extends into the IVS. This oligonucleotide was isolated by polyacrylamide gel electrophoresis and digested with either RNase T₂ or nuclease

Table 1 Molar yields of the modified nucleotides

Nucleotide species	Molar yield		
	Expected	Observed <i>SUP6</i>	<i>SUP6Δ32</i>
RNase A digests			
Gp	1	1	1
Cp	1	1.05	1.20
Up	3	3.18	4.25
ψp	1	0.90	0.08
mcm ⁵ s ² Up	1	0.80	0.92
ApCp	1	ND	ND
Api ⁶ ApApψp	1	ND	ND
RNase T ₂ digests			
Gp	1	1	1
Cp	2	2.17	2.22
Up	3	3.41	4.10
Ap	3	3.22	3.41
ψp	2	1.98	0.74
mcm ⁵ s ² Up	1	0.78	0.99
i ⁶ Ap	1	0.50	0.60

The amount of radioactivity in each nucleotide species shown in Fig. 5 was measured by cutting out the appropriate regions of the TLC plates and determining the c.p.m. by liquid scintillation counting. The observed molar yields were calculated relative to Gp, which is present once in the 13 base oligonucleotide. ND, not determined.

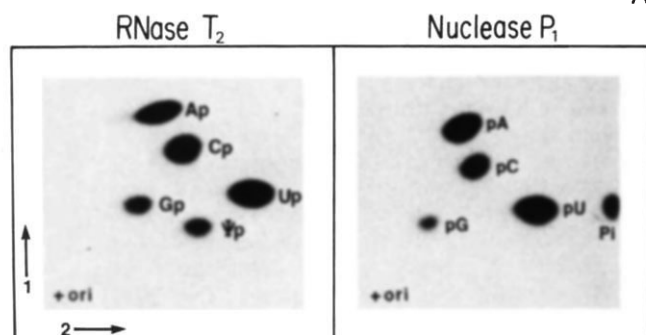


Fig. 6 Analysis of anticodon pseudouridine in pre-tRNA^{Tyr}. ³²P-labelled pre-tRNA^{Tyr} was isolated from strain M304 (carrying the *ts136* mutation at the *mal1* locus) as described by Ogden *et al.*⁴⁵. This RNA was digested with RNase T₁ (see legend to Fig. 4) and the products separated on a 20% polyacrylamide sequencing gel (not shown). The largest product (the 15 base anticodon oligonucleotide) was eluted from the gel and precipitated with 20 µg carrier tRNA. Half of the sample was digested with RNase T₂, and the remainder cleaved with nuclease P₁ (1 mg ml⁻¹ nuclease P₁ in 5 µl of 50 mM sodium acetate pH 5.0). The products were separated by two-dimensional TLC as in Fig. 5.

P₁, and the products separated by two-dimensional TLC. As seen in Fig. 6, RNase T₂ cleavage produces labelled ψp in equal amount to Gp, which is present once in the oligonucleotide. Nuclease P₁ generates mononucleotides with 5' phosphate and 3' hydroxyl ends and consequently leaves the 5' residue of a T₁ oligonucleotide unphosphorylated. Figure 6 shows that P₁ digestion produces no pψ, proving that the ψp in the RNase T₂ digest derives solely from the second position in the anticodon. Taken together, these experiments demonstrate that in the *in vivo* precursor tRNA^{Tyr}, uridine in the anticodon can be fully converted to ψ.

Discussion

We have established that deletion of the intron from a suppressor tRNA^{Tyr} gene causes: (1) a reduction in suppression; (2) a somewhat diminished amount of mature suppressor tRNA^{Tyr}; and (3) failure to produce a pseudouridine modification in the anticodon. As explained above, we believe that the reduced suppression seen in yeast harbouring *SUP6Δ32* cannot be solely accounted for by the decrease in the amount of tRNA produced by this gene. It seems likely that the suppression effect observed is due to the absence of ψ in the anticodon. There are precedents for a modification mutant that affects suppression, as a *Saccharomyces cerevisiae* mutant deficient in the synthesis of i⁶A has been shown to have a significantly decreased suppressor activity of another tyrosine-inserting ochre suppressor, *SUP7-1*³¹. Similar i⁶A-deficient mutants have been isolated in *Escherichia coli*³² and *Schizosaccharomyces pombe*³³. Whether the anticodon ψ affects the tRNA-ribosome interaction during translation or some other step such as the rate of aminoacylation is not known, although it should be possible to measure the latter by comparing modified and unmodified tRNA aminoacylation *in vitro*.

To date, we have not been able to detect any modification differences between *SUP6* and *SUP6Δ32* tRNA except at the anticodon position. However, our method of analysis has been limited to the 13 base suppressor specific oligonucleotide, and it is conceivable that modifications elsewhere in the tRNA molecule could be affected. Therefore we have attempted to isolate intact suppressor tRNA free of non-suppressor tRNA^{Tyr}

using the non-denaturing gel system developed by Korn³⁴. Although the suppressor tRNA isolated in this way is no more than 50% pure, we can detect no obvious differences in the modified bases from *SUP6* and *SUP6Δ32*, as analysed by two-dimensional TLC (data not shown). However, we cannot yet formally exclude the possibility that some other base modification is altered as a direct result of the lack of an intron, and that this defect in turn prevents the ψ modification from occurring.

There are two possible explanations for the dependence of the U to ψ on the presence of an IVS: (1) the IVS-containing precursor molecule is in fact the substrate for the appropriate modification enzyme, and the mature-sequence tRNA cannot be modified at this position; or (2) splicing is in some way a requisite step for the modification to occur (for example, by directing the tRNA to a subcellular location in which the modification activity resides). The finding that the anticodon ψ is present in high yield in the unspliced tRNA^{Tyr} precursor clearly favours the first hypothesis, although we do not know if naturally occurring tRNA precursors contain this modification (the precursor analysed having been produced in an *mal1* mutant strain). Ultimate proof of this model will require demonstration that a yeast cell-free extract can effect the U to ψ conversion in the precursor but not in the mature-sequence tRNA. We are currently attempting to answer this question using *in vitro* transcripts from *SUP6* and *SUP6Δ32* templates that are unmodified at this position. So far we have been able to detect ψ at position 39 (in the anticodon stem) but not at the anticodon in either the precursor or mature-sequence transcripts. This suggests that there may be a separate tRNA pseudouridine synthase for modification of the anticodon position. (There is evidence for multiple pseudouridylation enzymes in *Salmonella typhimurium*³⁵.) If this is the case in yeast, strains defective in the anticodon synthase activity may be present among collections of anti-suppressor mutants selected in tyrosine-inserting suppressor backgrounds^{31,36}.

Assuming that the IVS is necessary to create a structure capable of recognition by the anticodon ψ synthase, it would be of interest to know what features of the precursor structure contribute to this specificity. One obvious candidate is the (potentially) base paired region involving the anticodon and sequences within the IVS (Fig. 1), although in suppressor pre-tRNA^{Tyr} this structure may not exist, because only two of the three anticodon nucleotides can base pair. Site-directed mutagenesis of the IVS should allow identification of any role for base pairing in directing the modification. Such mutants should also be useful in delineating the precursor structural requirements for splicing.

tRNA^{Tyr} is the only yeast tRNA known to contain ψ in the anticodon, and the presence of introns in other tRNA species is not correlated with anticodon modifications in general. Therefore, whilst the results presented here point out a clear role for the intron in one tRNA gene family, a common function for all tRNA intervening sequences is not evident. Perhaps tRNA IVSs represent remnants of evolutionary gene rearrangements³⁷ and only occasionally evolve a role in RNA synthesis. Alternatively, there may be a common but as yet unidentified function for these IVSs, and the role for the IVS described here for tRNA^{Tyr} may represent an auxiliary use of the precursor RNA. Clearly, analysis of IVS mutants in other tRNA gene families will be necessary to obtain definitive answers to these questions.

We thank Rod Rothstein for stimulating discussions and for communicating results before publication, and Richard Ogden, Christine Guthrie, Marge Fidler, Ray Ng, Peter Gegenheimer and David Engelke for critical reading of the manuscript.

Received 9 December 1982; accepted 22 February 1983.

- Breathnach, R. & Chambon, P. *A. Rev. Biochem.* **50**, 349–383 (1981).
- Abelson, J. A. *Rev. Biochem.* **48**, 1035–1069 (1979).
- Lai, C.-J. & Khoury, G. *Proc. natn. Acad. Sci. U.S.A.* **76**, 71–75 (1979).
- Gruss, P., Lai, C. J., Dhar, R. & Khoury, G. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4317–4321 (1979).
- Hamer, D. H., Smith, K. D., Boyer, S. H. & Leder, P. *Cell* **17**, 725–735 (1979).

- Hamer, D. & Leder, P. *Cell* **18**, 1299–1302 (1979).
- Kedes, L. H. A. *Rev. Biochem.* **48**, 837–870 (1979).
- Alagstrom, P. *et al.* *Cell* **19**, 671–681 (1980).
- Nagata, S., Mantei, N. & Weissmann, C. *Nature* **287**, 401–408 (1980).
- Treisman, R., Novak, U., Favalaro, J. & Kamen, R. *Nature* **292**, 595–600 (1981).
- Gruss, P., Efstratiadis, A., Karathanasis, S., Konig, M. & Khoury, G. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6091–6095 (1981).

12. Ghosh, P. K. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 1386–1390 (1981).
13. *Mitochondrial Genes* (eds Slonimski, P., Borst, P. & Attardi, G.) (Cold Spring Harbor, New York, 1982).
14. Goodman, H. M., Olson, M. V. & Hall, B. D. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5453–5457 (1977).
15. Valenzuela, P., Venegas, A., Weinberg, F., Bishop, R. & Rutter, W. J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 190–194 (1978).
16. Venegas, A., Quiroga, M., Zaldivar, J., Rutter, W. J. & Valenzuela, P. *J. biol. Chem.* **254**, 12306–12309 (1979).
17. Kang, H. S., Ogden, R. C. & Abelson, J. in *Mobilization and Reassembly of Genetic Information* (eds Scott, W. A., Werner, R., Joseph, D. R. & Schultz, J.) 317–331 (Academic, New York, 1980).
18. Olson, M. V. *et al. Nature* **291**, 464–469 (1981).
19. O'Farrell, P. Z., Cordell, B., Valenzuela, P., Rutter, W. J. & Goodman, H. M. *Nature* **274**, 438–445 (1978).
20. Etcheverry, T., Colby, D. & Guthrie, C. *Cell* **18**, 11–26 (1979).
21. Peebles, C. L., Gegenheimer, P. & Abelson, J. *Cell* **32**, 525–536 (1983).
22. Munz, P., Amstutz, A., Kohli, J. & Leupold, U. *Nature* **300**, 225–231 (1982).
23. Wallace, R. B. *et al. Science* **209**, 1396–1400 (1978).
24. Scherer, S. & Davis, R. W. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4951–4955 (1979).
25. Orr-Weaver, T. L., Szostak, J. W. & Rothstein, R. J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6354–6358 (1981).
26. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
27. Olson, M. V. *et al. Nature* **267**, 639–641 (1977).
28. Saneyoshi, M., Ohashi, Z., Harada, F. & Nishimura, S. *Biochim. biophys. Acta* **262**, 1–10 (1972).
29. Colby, D. *et al.* (in preparation).
30. Kang, H. S., Ogden, R. C., Knapp, G., Peebles, C. L. & Abelson, J. in *Eukaryotic Gene Regulation* (eds Axel, R., Maniatis, T. & Fox, C. F.) 69–83 (Academic, New York, 1979).
31. Laten, H., Gorman, J. & Bock, R. M. *Nucleic Acids Res.* **5**, 4329–4342 (1978).
32. Eisenberg, S. P., Soll, L., & Yarus, M. in *Transfer RNA: Biological Aspects* (eds Soll, D., Abelson, J. & Schimmel, P.) 469–480 (Cold Spring Harbor, New York, 1980).
33. Janner, F., Vogeli, G. & Fluri, R. *J. molec. Biol.* **139**, 207–219 (1980).
34. Korn, L. J. & Gurdon, J. B. *Nature* **289**, 461–465 (1981).
35. Green, C. J., Kammen, H. O. & Penhoet, E. E. *J. biol. Chem.* **257**, 3045–3052 (1982).
36. McCready, S. J. & Cox, B. S. *Molec. gen. Genet.* **124**, 305–320 (1973).
37. Gilbert, W. *Nature* **271**, 501 (1978).
38. *Transfer RNA: Structure, Properties and Recognition* (eds Schimmel, P. R., Soll, D. & Abelson, J. N.) 518–519 (Cold Spring Harbor Laboratory, New York, 1979).
39. Hsiao, C.-L. & Carbon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3829–3833 (1979).
40. Davis, R. W. *et al. Meth. Enzym.* **65**, 404–411 (1980).
41. Thomas, P. S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201–5205 (1980).
42. Sherman, F., Fink, G. R. & Lawrence, C. W. in *Methods in Yeast Genetics*, 61–64 (Cold Spring Harbor, New York, 1979).
43. Rubin, G. M. *Meth. Cell Biol.* **12**, 45–64 (1975).
44. Knapp, G., Beckmann, J. S., Johnson, P. F., Fuhrman, S. A. & Abelson, J. *Cell* **14**, 221–236 (1978).
45. Ogden, R. C. *et al. Cell* **17**, 399–406 (1979).

LETTERS TO NATURE

A proof of the impossibility of inductive probability

Karl Popper

Penn, High Wycombe, Buckinghamshire, UK

David Miller

Department of Philosophy, University of Warwick,
Coventry CV4 7AL, UK

Proofs of the impossibility of induction have been falling 'dead-born from the Press' ever since the first of them (in David Hume's *Treatise of Human Nature*) appeared in 1739. One of us (K.P.) has been producing them for more than 50 years. The present proof strikes us both as pretty.

Let h be a hypothesis, and e (possible) evidence in favour of it. In a simple case e is deducible from h in the presence of background knowledge b (b includes the initial conditions needed to derive the predictions e from h , and it may for the time being be regarded as unproblematic). Thus we have

$$p(e, hb) = 1 \quad (1)$$

which can be read: the probability of e , given h , in the presence of b , equals 1. We also have

$$p(he, b) = p(h, b) \quad (2)$$

since e is logically contained in h (in the presence of b). On the other hand we have the product law,

$$p(he, b) = p(h, eb)p(e, b) \quad (3)$$

which means that, provided

$$0 < p(e, b) < 1 \quad (4)$$

we obtain from equations (2)–(4)

$$p(h, eb) > p(h, b) \quad (5)$$

provided $p(h, b) > 0$.

Now this explains, and seems to justify, the belief in induction. For if (in the presence of b) e follows from h , then the probability of h , given e , is greater than the prior probability of h , before e was given. So we may say that e supports h .

It is the support of h by e , the increase in the probability of h generated by the evidence e , that looks like a probabilistic inductive effect. We now show that this is an illusion.

Any proposition, for example, the hypothesis h , can be factorized with respect to any other proposition, for example, the empirical evidence statement e , into two factors:

$$h = (h \leftarrow e)(h \vee e) \quad (6)$$

where $h \leftarrow e$ (read ' h if e ') is the same as $e \rightarrow h$ (read ' e then h ').

This factorization of h has important properties, (i) and (ii).

(i) Any weakening of the content of either of the two factors makes the right-hand side of equation (6) weaker than the left-hand side. (A proposition x is logically weaker than a proposition y if, and only if, x follows from y but not the other way round.)

(ii) Any proposition x that can replace one of the two factors in the equation is logically at least as strong as the factor that it replaces: the factor follows deductively from x .

Thus each of the two factors is the weakest proposition strong enough in the presence of the other factor to entail the proposition h .

Given the evidence e the probability of the second factor, $h \vee e$, equals 1 (since it follows from e); moreover, given e , the probability of the first factor, $h \leftarrow e$, equals the probability of h , given e : for

$$p(h \leftarrow e, e) = p((h \leftarrow e)e, e) = p(he, e) = p(h, e) \quad (7)$$

Thus

$$\begin{aligned} p((h \leftarrow e)(h \vee e), e) &= p(h, e) = p(h \leftarrow e, e) \\ &= p(h \leftarrow e, e)p(h \vee e, e) \end{aligned} \quad (8)$$

that is: given e , the factors of h are probabilistically independent: the probability of their product is equal to the product of their probabilities.

Thus we have split h into two factors, the second of which, $h \vee e$, deductively follows from the evidence e : it is the logically strongest part of h (or of the content of h) that follows from e . The first factor, $h \leftarrow e$, contains all of h that goes beyond e .

In the presence of e , the first factor, $h \leftarrow e$, is deductively equivalent to h itself. (All this, from equation (7) on, holds also if we assume the presence of the background knowledge b .)

To sum up: we have split h with respect to e into two factors. One of them follows deductively from e , and it can be, in the presence of e (and of b), ignored; the other is, in the presence of e (or of e and b), equivalent to h .

The whole question of induction can now be put as follows: Let us assume that e supports h ; that is, the probability of

h , given e , is greater than its probability without e :

$$p(h, e) > p(h)$$

or, making the background knowledge explicit,

$$p(h, eb) > p(h, b)$$

Does e , in this case, provide any support for the factor $h \leftarrow e$, which in the presence of e is alone needed to obtain h ? The answer is: No. It never does. Indeed e countersupports $h \leftarrow e$ unless either $p(h, e) = 1$ or $p(e) = 1$ (which are possibilities of no interest).

This can be shown by:

Theorem 1: If

$$p(h, e) \neq 1 \neq p(e)$$

then

$$p(h \leftarrow e, e) < p(h \leftarrow e)$$

This means: e countersupports $h \leftarrow e$.

Our theorem can be strengthened: the degree of countersupport equals $Exc(h, e)$, that is the excess of the probability of the conditional over the conditional probability.

Theorem 2: Under the same assumptions,

$$\begin{aligned} p(h \leftarrow e) - p(h \leftarrow e, e) &= p(h \leftarrow e) - p(h, e) \\ &= p(-h, e)p(-e) = Exc(h, e) > 0 \end{aligned}$$

This is again unaffected if we bring in the background knowledge b , as shown by

Corollary: If

$$p(h, eb) \neq 1 \neq p(e, b)$$

then

$$\begin{aligned} p(h \leftarrow e, b) - p(h \leftarrow e, eb) &= p(h \leftarrow e, b) - p(h, eb) \\ &= p(-h, eb)p(-e, b) = Exc(h, e, b) > 0 \end{aligned}$$

The proof of this corollary, of course, proves also theorem 1 and theorem 2.

Proof: We transform $p(-h, eb)p(-e, b)$ so as to obtain the differences on the left-hand side of the corollary.

$$\begin{aligned} p(-h, eb)p(-e, b) &= (1 - p(h, eb))(1 - p(e, b)) \\ &= 1 - p(h, eb) - p(e, b) + p(h, eb)p(e, b) \\ &= 1 - p(h, eb) - p(e, b) + p(h, eb) \quad (\text{product law}) \\ &= 1 - (p(e, b) - p(h, eb)) - p(h, eb) \\ &= (1 - p(-h, eb)) - p(h, eb) \\ &= p(h \leftarrow e, b) - p(h, eb) \end{aligned}$$

In view of equation (7) this completes the proof.

All this means: that factor that contains all of h that does not follow deductively from e is strongly countersupported by e . It is countersupported the more the greater the content of e , which may be measured by

$$ct(e, b) = 1 - p(e, b) \quad (9)$$

Indeed, the countersupport increases with the content of e , whether e supports h or not.

This result is completely devastating to the inductive interpretation of the calculus of probability. All probabilistic support is purely deductive: that part of a hypothesis that is not deductively entailed by the evidence is always strongly countersupported by the evidence—the more strongly the more the evidence asserts. This is completely general; it holds for every hypothesis h ; and it holds for every evidence e , whether it supports h , is independent of h , or countersupports h .

There is such a thing as probabilistic support; there might even be such a thing as inductive support (though we hardly think so). But the calculus of probability reveals that probabilistic support cannot be inductive support.

High-resolution X-ray and radio maps of the millisecond pulsar

R. H. Becker

Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061, USA

D. J. Helfand

Columbia Astrophysics Laboratory, Columbia University, New York, New York 10027, USA

The detection¹ of a new pulsar in the radio source 4C21.53 with a period of 1.5 ms has raised important questions concerning the origin and evolution of galactic neutron stars. Before its discovery, we had been conducting radio and X-ray observations of 4C21.53 designed to place in context the intriguing properties of the source (extended emission, a low-frequency excess, and interstellar scintillation²). We present here the results of this programme including a stringent upper limit on any X-ray emission from the pulsar and high-resolution multi-frequency radio maps of the region which provide the best available pulsar position. We argue against any connection between the pulsar and the nearby extended radio emission region and comment on the constraints which the X-ray data place on the nature of this object.

The radio images of the region were obtained with the Very Large Array (VLA) during two observing runs. In July 1982, the region was observed at 20- and 6-cm wavelength in the 'B' configuration, resulting in maps at the two wavelengths with angular resolution of ~ 2.6 and ~ 0.8 arc s and fields of view of 30 and 10 arc min, respectively. In December 1982, the same region was observed in the 'D' configuration at 2-cm wavelength with angular resolution and field size of ~ 3 arc s and 3 arc min, respectively. The 20-cm wavelength image is shown in Fig. 1. Three sources are easily discernible: an extended source, 4C21.53W, and two unresolved objects. The position and fluxes of all three are listed in Table 1. The southern compact object is the new fast pulsar.

In our 6-cm map of the region, the compact source to the east of 4C21.53W was detected with a flux of 2 ± 0.4 mJy. The fast pulsar was not apparent in the data and we can place a 2σ upper limit to its 6 cm flux of 0.4 mJy. This requires that the pulsar spectral index of ~ 2.5 between 609 and 1,415 MHz (ref. 1) steepens to ~ 3.5 above 1,415 MHz. Such spectral breaks are seen in several pulsars, although few have quite such a steep index. The Crab pulsar has a spectral slope of just 3.5 at 1,400 MHz.

The extended source can be described as a centrally peaked, flat-spectrum source of emission, a description which is equally appropriate for either an H II region of a Crab-like supernova remnant (SNR). This ambiguity can be broken by either the measurement of hydrogen recombination lines (indicative of H II regions) or linear polarization (indicative of Crab-like SNR). The VLA observations sample all four Stokes parameters of the radio radiation, allowing a measurement of the degree of polarization. For example, PSR1937+214 has an integrated polarization at 20 cm of $10 \pm 2\%$. However, we have found no evidence of polarization in 4C21.53W at any of the three wavelengths observed, with upper limits at 20-, 6- and 2-cm wavelengths of 1%, 3% and 5%, respectively. Contrary to early reports, there is no evidence in the morphology of this region for a connection to the pulsar, and 21-cm absorption profiles for both the pulsar and the extended region have now been obtained³ which show that the two must be at different distances. We conclude that 4C21.53W is an unrelated H II region whose proximity to the pulsar is important only in that it provided the

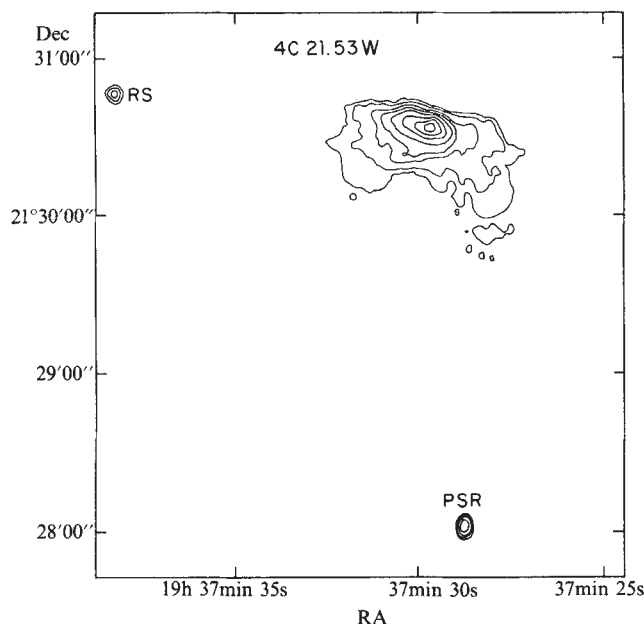


Fig. 1 A 20-cm map of 4C21.53W obtained with the VLA B-array in 1982 July. The peak flux is 27 mJy per beam and the contour levels are 0.03, 0.05, 0.10, 0.15, 0.25, 0.45, 0.65 and 0.85 of this value. The background radio source (RS) and millisecond pulsar PSR1937+214 (PSR) are labelled.

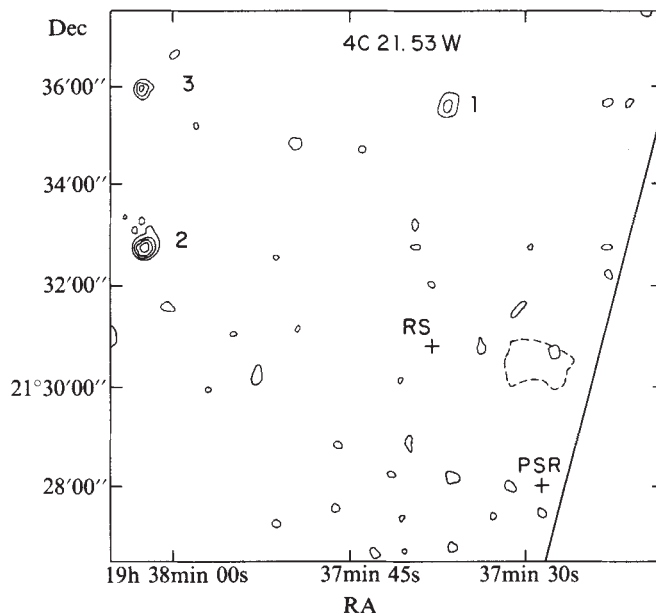


Fig. 2 An Einstein high-resolution imager X-ray (0.1–3.0 keV) map of the vicinity of 4C21.53W. Contour levels are 5, 10, 15 and 20×10^{-3} counts arc min $^{-2}$ s $^{-1}$ above a uniform, subtracted background level of 4.5×10^{-3} counts arc min $^{-2}$ s $^{-1}$. The positions of the background radio source and PSR1937+214 are noted; the dashed contour is the lowest contour from the map of the H II region in Fig. 1. The diagonal line at the right is the edge of the instrument field of view. Parameters of the three detected X-ray sources, together with upper limits for the radio objects, are given in Table 1.

stimulus for much of the work leading to the discovery of PSR1937+214.

The region in the vicinity of 4C21.53 was also observed on 10–11 April 1981 with the high resolution imager (HRI has 4 arc s resolution, 0.1–3.5 keV) on board the Einstein Observatory⁴. A total of 15,300 s of good data were collected; a map is shown in Fig. 2. Three X-ray sources were detected at greater than 3σ above the mean field background rate of 5.5×10^{-3} counts arc min $^{-1}$ s $^{-1}$ and are labelled in Fig. 2. The positions and count rates for these sources are given in Table 1b. (A reasonable count rate to flux conversion factor is 10^{-10} erg cm $^{-2}$ s $^{-1}$ counts $^{-1}$.) Also noted on Fig. 2 are the positions of the serendipitous radio source (RS) seen in the VLA map and PSR1937+214; the lowest contour of the H II region from Fig. 1 is shown as a dashed line. No detectable X-ray emission was associated with these three objects. The background-subtracted X-ray counting rates within circles of radius 12 arc s centred on the radio point source positions are $0.7 \pm 3 \times 10^{-4}$ counts s $^{-1}$ and $3 \pm 3 \times 10^{-4}$ counts s $^{-1}$ for the unidentified radio source and the pulsar, respectively. The 2σ upper limit to emission from the region of the extended radio source is 3×10^{-3} counts s $^{-1}$.

Four types of X-ray emission are known to be associated with radio pulsars: extended thermal emission from the super-

nova remnant in which the pulsar was created (such as Vela and MSH15–52)⁵, extended nonthermal emission generated by synchrotron electrons accelerated by the pulsar (such as the Crab Nebula and several older pulsars)⁶, nonthermal magnetospheric emission (such as the Crab pulsar) and thermal emission from the stellar surface, either from the initial cooling of the star or as a result of internal heating mechanisms (perhaps, PSR1929+10)^{6,7}. The X-ray map of 4C21.53 sets limits on each of these mechanisms for PSR1937+214. For a distance¹ of 2.5 kpc, the HRI limit corresponds to an X-ray luminosity limit of $\leq 7 \times 10^{31}$ erg s $^{-1}$ for a Crab-like spectrum modified by absorption of a mean interstellar density of 1 cm $^{-3}$. The corresponding blackbody temperature limit for a star of radius 15 km is $\sim 1.5 \times 10^6$ K.

The lack of any radio or X-ray emission from a SNR in the vicinity of PSR1937+214 suggests either that the object is $\gg 10^4$ yr old or that it formed without leaving a canonical SNR. The most recent value for P of 1.4×10^{-19} s $^{-1}$ (which is subject to a revision downward by $\sim 30\%$ when the source position reported here is used in the timing solution)³ implies a characteristic age of $> 10^8$ yr, consistent with the former hypothesis.

Table 1 Sources in the vicinity of the millisecond pulsar

Name	RA (1950)	Dec (1950)	Flux (mJy)
a Radio sources			
RS	19 h 37 min 38.22 s $\pm$ 0.02	+21° 30' 46.7" $\pm$ 0.2	5.5 $\pm$ 0.5 (20 cm) 2.5 $\pm$ 0.4 (6 cm)
PSR1937+214	19 37 28.74 $\pm$ 0.02	+21 28 01.8 $\pm$ 0.2	14 $\pm$ 1 (20 cm) <0.4 $\pm$ 0.4 (6 cm) Flux (10^{-3} HRI counts s $^{-1}$)
b X-ray sources			
1	19 37 36.63 $\pm$ 0.25	+21 35 31.8 $\pm$ 3.5	1.4 $\pm$ 0.4
2	19 38 02.50 $\pm$ 0.25	+21 32 48.2 $\pm$ 3.5	2.5 $\pm$ 0.5
3	19 38 02.74 $\pm$ 0.25	+21 35 56.5 $\pm$ 3.5	1.4 $\pm$ 0.4
RS	As above	As above	<0.7 (2σ)
PSR1927+214	As above	As above	<1.0 (2σ)

Optical and X-ray magnetospheric emission is expected to scale as $P^{-8}P^2$ (ref. 8); comparison to the Crab's high-energy pulsed emission would then suggest a time-averaged optical magnitude of $m_v \sim 22$ and an X-ray flux of ~ 0.1 HRI counts s^{-1} . While the published optical data⁹ do not rule out such pulsations, our X-ray limit is considerably below this prediction, indicative of the substantial differences in the magnetospheric structure of PSR1937+214 (light cylinder radius $R_{LC} \sim 5R_*$) compared with that of other radio pulsars (for which $R_{LC} > 10^2 R_*$). Significant γ -ray emission is thus also unlikely.

The rotational energy loss rate for the neutron star, $\dot{E} \sim 10^{36}$ erg s^{-1} . Our observations require that $< 5 \times 10^{-5}$ of this energy loss appear as soft X-rays. This is the lowest such efficiency limit for any of the ~ 30 pulsars observed with Einstein⁶ and can be used to set limits on some of the pulsar's characteristics relevant to either the formation of a synchrotron nebula or the internal heating of a neutron star. For example, any synchrotron nebula around this pulsar is less luminous than that produced around the short period source PSR1055-52 (ref. 10), which has an energy loss rate of only a few per cent of that for PSR1937+214. This is probably attributable to a combination of factors including a much weaker pulsar magnetic

field and a magnetosphere too small to accommodate the required acceleration of particles to energies of 10^{14} eV. The lack of polar cap heating from backflowing particles is also the result of the restricted magnetosphere, while the low level of surface thermal emission is expected for a star which is both old and has a low magnetic field.

A spectrum of PSR1937+214 including all available radio, IR, optical and X-ray upper limits shows that, below 10^{19} Hz, $< 10^{-3}$ of the spin down energy emerges as electromagnetic radiation. As noted above, future more stringent limits on the source's γ -ray emission is unlikely significantly to alter this conclusion. The lack of an optical or radio synchrotron nebula surrounding the object suggest that the fraction of $\dot{E}$ which emerges as particles is also small. Development of a magnetospheric model for this source as well as future timing observations to limit the star's precession amplitude are needed before the relative contributions to $\dot{E}$ of magnetic dipole and gravitational quadrupole radiation can be assessed.

We acknowledge the support of NASA under contracts NAS 8-30753 and NAS 8-432 and the Air Force under grant AFOSR 82-0014. This is Columbia Astrophysics Laboratory Contribution no. 237.

Received 23 February; accepted 14 March 1983.

1. Backer, D., Kulkarni, S., Heiles, C., Davis, M. & Goss, M. *Nature* **300**, 615-618 (1982).
2. Erickson, W. C. *Astrophys. J. Lett.* **264**, L13-L17 (1983).
3. Backer, D. C. *et al. Bull. Am. astr. Soc.* (in the Press).
4. Giacconi, R. *et al. Astrophys. J.* **230**, 540-550 (1979).

5. Seward, F. *Proc. IAU Symp.* No. 101 (1983).
6. Helfand, D. *Proc. IAU Symp.* No. 101 (1983).
7. Helfand, D., Chanan, G. & Novick, R. *Nature* **283**, 337-343 (1980).
8. Pacini, F. *Astrophys. J. Lett.* **163**, L17-L19 (1970).
9. Djorgovski, S. *Nature* **300**, 618 (1982).
10. Cheng, A. & Helfand, D. *Astrophys. J.* (in the press).

Polarimetry of the millisecond pulsar

Daniel R. Stinebring

National Radio Astronomy Observatory, Edgemont Road,
Charlottesville, Virginia 22901, USA

The millisecond pulsar¹ PSR1937+21 might be expected to have unique waveform and polarization properties as a result of its remarkably short period and small spin-down rate^{2,3}. The proximity of the velocity of light cylinder to the star ($R_{LC} = 7.4 \times 10^6$ cm) and the low value of inferred magnetic field ($\sim 10^8$ G, 10^{-4} that of the Crab pulsar) should strongly influence the magnetospheric structure and particle density of this pulsar. And yet the waveform is similar to that of the Crab pulsar⁴ and PSR1055-52 (ref. 5), which have interpulses that are comparable in strength to the main pulse, and the polarization properties are at first sight unexceptional². I present here polarization observations of the main pulse and interpulse at 1,415 and 2,380 MHz that have higher time resolution than previously available. The main pulse depolarizes substantially over this frequency range whereas the interpulse polarization almost doubles. There is evidence that orthogonally polarized radiation is present on the leading edge of the main pulse and the trailing edge of the interpulse and that it accounts for the $\sim 90^\circ$ difference in position angle between the main pulse and interpulse. The interpulse decreases in intensity relative to the main pulse as f^{-1} , indicating (if the trend continues to lower frequency) that the high frequency interpulse dominates the main pulse below ~ 700 MHz. The main pulse is found to consist of two closely spaced components.

The observations were made between 23 January and 30 January 1983 with the 305-m telescope at the Arecibo Observatory. Synchronous averages of the polarization data were made on-line, with 100 and 200 sample points across the pulsar period at 1,415 and 2,380 MHz, respectively. Bandpass filters of 125 and 250 kHz were used at 1,415 and 2,380 MHz, which, with RC time constants of 20 and 10 μs , resulted in effective time constants of 30 and 14 μs when dispersion smearing was accounted for. Standard adding polarimetry techniques⁶ were used on the circularly polarized inputs, although circular to

linear cross-coupling was not corrected for because of the small intrinsic circular polarization found for this pulsar.

Figure 1 presents the 1,415 MHz polarization results and shows the following features: an unresolved main pulse and interpulse separated by 173° of longitude; moderately strong linear polarization across the main pulse and interpulse with significant depolarization on the inner edges of the pulses (the edges facing across the 173° separation); a position angle separation of close to 90° between main pulse and interpulse; and an essential absence of circular polarization (the 3% features visible are statistically significant but are not reliable because of $\sim 10\%$ uncorrected circular to linear cross-coupling in the feed antenna). Ashworth *et al.*² have presented polarization observations of comparable quality made at 1,667 MHz. Except for minor differences in relative intensities, which are consistent with the spectral development between 1,415 and 2,380 MHz, there is very close agreement between the two data sets. Table 1 summarizes the major observational properties of this pulsar and includes values obtained from their paper.

The 2,380-MHz data displayed in Fig. 2, with time resolution about twice as good as at 1,415 MHz, show a shoulder on the trailing edge of the main pulse. The expanded version of the main pulse shown in Fig. 3 reveals that the pulse consists of two components, the weaker component occurring about 35 μs or 8° of longitude after the main pulse peak and rising to $\sim 40\%$ of the leading component peak intensity. The integrated linear polarization has decreased to a third of its 1,415 MHz value in the main pulse, and there is evidence in Fig. 3 that part of this depolarization is due to the superposition of orthogonal polarization states from the two main pulse components. As at 1,415 MHz, the inner edge of the main pulse is also significantly depolarized. The interpulse, on the other hand, shows more than a 50% increase in percentage linear intensity over the 1,415 MHz value. This occurs partly through an overall decrease in the unpolarized interpulse intensity relative to the main pulse and partly through the restoration of linear polarization on the trailing edge of the interpulse. The position angle difference between the main pulse and interpulse remains close to 90° at 2,380 MHz. The circular polarization again is present at too low a level to be reliable.

The main pulse and interpulse are significantly wider at 1,415 MHz than at 2,380 MHz (after correction for instru-

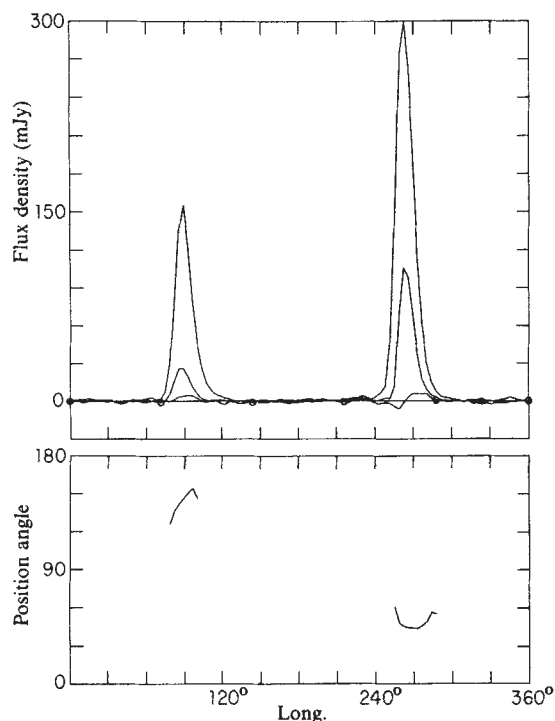


Fig. 1 Polarization waveform of PSR1937+21 at 1,415 MHz. The total intensity (outer envelope), linear intensity (middle envelope), and circular intensity (circled, near zero) are shown for an average of 6.4×10^6 pulses. The effective time constant is 7.1° of longitude or $30 \mu\text{s}$ and there are 100 samples across the period. Position angles are referred to an arbitrary origin and are only plotted when the linear intensity is greater than two times the r.m.s. noise level. The flux density scale is an average over many scintillation bands.

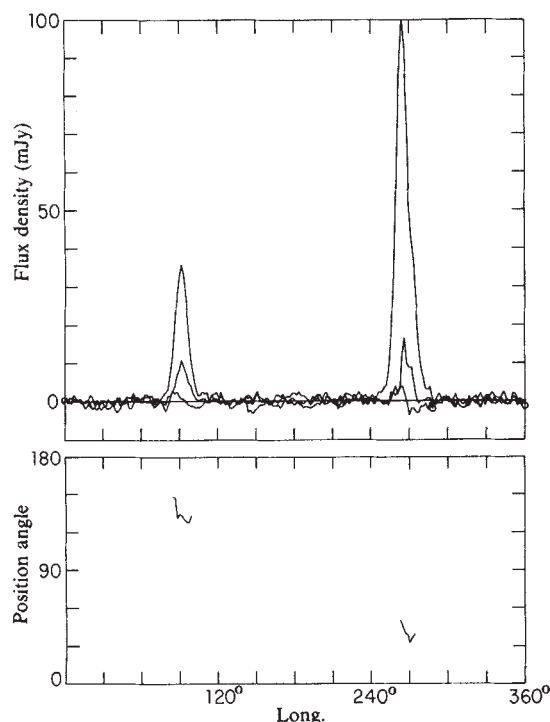


Fig. 2 Polarization waveform at 2,380 MHz for an average of 3.1×10^6 pulses. The total intensity (outer envelope), linear intensity (middle envelope) and circular intensity (circled, near zero) are shown. The effective time constant is 3.2° of longitude or $14 \mu\text{s}$ and there are 200 samples across the period. The sign of the circular polarization and the sense of position angle rotation may be reversed from that in Fig. 1. Position angles are referred to an arbitrary origin and the flux density scale is approximate.

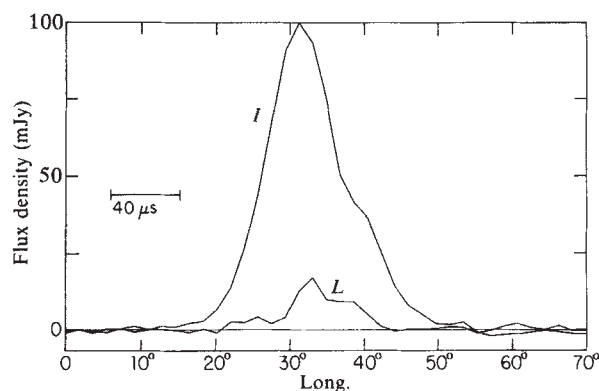


Fig. 3 Expanded plot of the total intensity and linear polarization for the main pulse at 2,380 MHz (same time constant and sample interval as in Fig. 2). A second emission component is visible on the trailing edge of the pulse. The longitude axis is shifted with respect to Figs 1 and 2.

mental broadening), although Ashworth *et al.*² report values consistent with the 2,380-MHz widths. If the broadening is real it may be due, at least in the main pulse, to changing relative intensities of the two emission components.

Explanations for interpulse emission⁷⁻¹⁰ divide into 'one-pole' models, in which both main pulse and interpulse are produced above a single magnetic pole, and 'two pole' models, where the main pulse and interpulse arise from opposite magnetic poles. Ashworth *et al.*² note that the 90° position angle difference between the main pulse and interpulse is hard to reconcile with a two-pole model since near equality of the position angles would be expected. There is no conflict with a two-pole model, however, if the 90° position angle difference is due to the orthogonally polarized radiation phenomenon seen in many other pulsars and hinted at here by the depolarization on the inside edges of the pulses. The sign of position angle rotation across the two pulses is expected to be identical in sense for most one-pole models and opposite in sense for most two-pole models. Although the 2,380-MHz data in Fig. 3 support a one-pole interpretation, the 1,415-MHz position angle does not rotate monotonically across either the interpulse or the main pulse and some sub-averages of the 2,380 MHz data show an interpulse position angle turnover at 1,415 MHz. Higher sensitivity, higher time-resolution data are needed to clarify the position angle behaviour.

Although the main pulse spectral index ($S \propto f^\alpha$; $\alpha = -2.0 \pm 0.5$) is poorly determined because of strong scintillation, the spectral index difference between the main pulse and the interpulse is close to 1.0 over this frequency range. If this behaviour continues to lower frequency, the presently designated 'interpulse' becomes equal to the main pulse near $f_{eq} \sim 700$ MHz and would have been designated the main pulse in a lower-frequency discovery of the pulsar. The Crab pulsar has a similar spectral index difference between the main pulse and the

Table 1 Main pulse (MP) and interpulse (IP) properties

	1,415 MHz	1,667 MHz*	2,380 MHz
Energy (IP)/energy (MP)	0.50	0.40	0.33
MP to IP separation (deg)	173 ± 1.8	173	173 ± 0.9
MP width† (μs)	65 ± 10	40 ± 10	43 ± 4
IP width† (μs)	56 ± 10	$\leq 40 \pm 10$	47 ± 4
L/I (MP)‡	0.30	0.29	0.11
L/I (IP)‡	0.16	0.20	0.29
$PA(IP) - PA(MP)$ §	$100^\circ \pm 4^\circ$	$\sim 90^\circ$	$90^\circ \pm 8^\circ$

* From ref. 2.

† Full-width half-intensity, corrected for instrumental broadening.

‡ Stokes parameter: L , linear intensity; I , total intensity.

§ Position angle measured at the linear polarization peaks.

interpulse, which results in near equality of the two pulses at $f_{eq} \sim 200$ MHz (ref. 11).

Although the waveform and polarization properties of the millisecond pulsar are qualitatively similar to the Crab pulsar, it is not yet clear whether this is important or simply a coincidence arising from the limited frequency range over which the millisecond pulsar has been studied. If the similarity is borne out by further observations, particularly at lower frequencies, it will be challenging to understand how the pulsar mechanism can produce comparable results in such disparate physical settings.

I thank R. Murphy, P. M. B. Shames, M. M. Davis and W. R. Stinebring for help with the data-taking program and other assistance. The National Radio Astronomy Observatory is operated by Associated Universities, Inc., under contract with the NSF. The Arecibo Observatory is operated by Cornell University under contract with the NSF.

Received 28 February; accepted 14 March 1983.

1. Backer, D. C., Kulkarni, S. R., Heiles, C., Davis, M. M. & Goss, W. M. *Nature* **300**, 615-618 (1982).
2. Ashworth, M., Lyne, A. G. & Smith, F. G. *Nature* **301**, 313-314 (1983).
3. Backer, D. C., Kulkarni, S. R. & Taylor, J. H. *Nature* **301**, 314-315 (1983).
4. Manchester, R. N. *Astrophys. J. Lett.* **163**, L61-L63 (1971).
5. McCulloch, P. M., Hamilton, P. A., Ables, J. G. & Komesaroff, M. M. *Mon. Not. R. astr. Soc.* **175**, 71P-75P (1976).
6. Stinebring, D. R., Cordes, J. M., Rankin, J. M., Weisberg, J. M. & Boriakoff, V. *Astrophys. J.* (submitted).
7. Manchester, R. N. & Lyne, A. G. *Mon. Not. R. astr. Soc.* **181**, 761-767 (1977).
8. Cheng, A. F. & Ruderman, M. A. *Astrophys. J.* **216**, 865-872 (1977).
9. Hankins, T. H. & Cordes, J. M. *Astrophys. J.* **249**, 241-253 (1981).
10. Narayan, R. & Vivekanand, M. *Nature* (submitted).
11. Rankin, J. M. *et al. Astrophys. J.* **162**, 707-725 (1970).

Global distribution and southern hemispheric trends of atmospheric CCl_3F

P. J. Fraser*, P. Hyson, I. G. Enting & G. I. Pearman

CSIRO Division of Atmospheric Physics, Aspendale, Victoria, Australia 3195

The spatial and temporal variability of the tropospherically inert tracer, trichlorofluoromethane (CCl_3F), has been simulated using a global atmospheric transport model, incorporating an advective-diffusive transport scheme and known release and photolytic data. The observational data are taken from the Geophysical Monitoring for Climatic Change (GMCC) global network^{1,2}, from the Pacific north-west (PNW) USA and South Pole³, and from the CSIRO southern hemispheric stations at Cape Grim, Tasmania, and Mawson, Antarctica^{4,5}. The resulting global CCl_3F distribution is shown in Fig. 1. A current CCl_3F atmospheric lifetime of 75 yr is obtained. The observations suggest that small, residual errors may exist in the CCl_3F release data.

Cumulative CCl_3F emissions including 1979 are 2.1% higher than used in a previous study⁶, based on a revised estimate of global releases up to 1980⁶, incorporating previously unaccounted for CCl_3F losses at production sites (1.5% increase in total emissions over the period 1935-79), enhanced refrigeration losses (0.3%, 1937-79) and expanded USSR-Eastern European sources (0.3%, 1975-79). The model is initiated in 1959 and cumulative, pre-1960 emissions, reduced by 4% for photolytic losses, are released in 1959. All CCl_3F concentrations reported here are in a scale established by R. A. Rasmussen (Oregon Graduate Center) multiplied by 0.952, derived from inter-laboratory calibration experiments (Fig. 2). The linearity of these results suggests that Rasmussen's calibrations scale has not changed with time.

The two-dimensional model⁵ simulates global atmospheric transport primarily by advection, using observed wind fields,

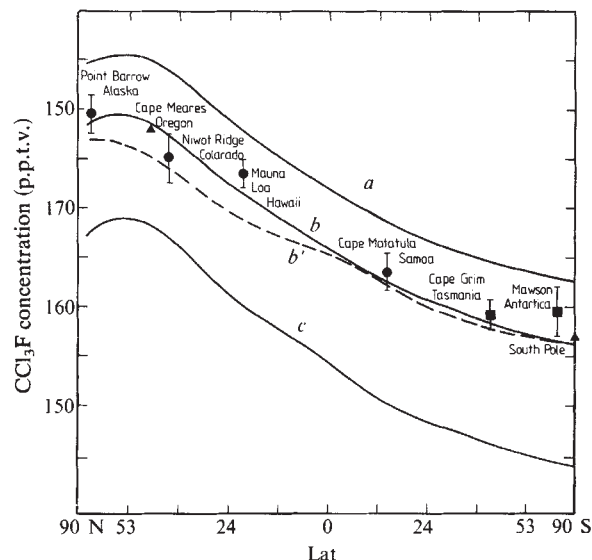


Fig. 1 Observed and calculated meridional distributions of CCl_3F for mid-January 1980. Data sources: ●, refs 1, 2; ▲, ref. 3; ■, this work. Error bars are standard errors of linear regression estimates. Point Barrow, Cape Meares (PNW, USA), Cape Matatula, Cape Grim and Mawson are all sea-level stations, Niwot Ridge is at 3.7 km, Mauna Loa 3.4 km and South Pole 2.8 km. Curves a, b and c are model calculations (1,000 mbar) for three values of K_{zz} in the stratosphere corresponding to τ_c values of 100, 75 and 50 yr respectively. Curve b' is the same as b for 700 mbar.

and also by empirically derived eddy diffusion, and incorporates CCl_3F release as a function of time and latitude. The parameterization of stratospheric tracer transport has been revised because, for quasi-inert tracers, it has been found that off-axis diffusive coefficients (K_{yz} , K_{zy}) are approximately equal and of opposite sign^{7,8}. In these conditions the effect of eulerian advection, as modelled by meridional and vertical winds, is approximately cancelled by that of eddy diffusion, whose major component is represented by these off-axis terms^{7,8}. General circulation models show that stratospheric ozone transport by mean meridional circulation is largely cancelled by that due to zonal eddies⁹. To approximate this cancellation in the model, stratospheric transport was simulated using only the minor, on-axis diffusive coefficients (K_{yy} and K_{zz}). This modification allows a reasonable simulation of the photolytic sink for CCl_3F in the stratosphere, using published photodissociation coefficients¹⁰ and a radiation scheme including latitudinal variation of day length and solar zenith angle¹¹. Following this major change in the method of simulating stratospheric transport, the horizontal diffusive coefficients (K_{yy}) and the stratospheric vertical coefficients (K_{zz}) used previously⁵ were increased by 15 and 80% respectively to refine agreement between simulations and observations (Figs 1 and 3).

Other two-dimensional models with fixed vertical transport overestimate stratospheric CCl_3F in both the tropics and mid-latitudes¹², presumably because photolysis rates are underestimated. Enhanced penetration of UV radiation through the Shumann-Runge region of the molecular oxygen spectrum¹³ will yield faster CCl_3F photolysis rates and perhaps explain the above discrepancies. In our model we use published photolysis rates¹⁰ and force agreement with observations by adjusting vertical transport rates.

The globally averaged CCl_3F concentration for mid-January 1980 (Fig. 1) was 168 parts per 10^{12} by volume (p.p.t.v.), compared with the model simulation, using concentrations from appropriate model locations, of 166 p.p.t.v. Both observations and model now indicate an interhemispheric difference of 10% (mid-January 1980) compared with 12% from the earlier model⁵. This transport scheme gives a current atmospheric lifetime (τ_c = current atmospheric content/current destruction rate) of 75 yr. Assuming 5% uncertainties in CCl_3F release

* Present address: Cooperative Institute for Research in Environmental Sciences, University of Colorado/NOAA, Boulder, Colorado 80303, USA.

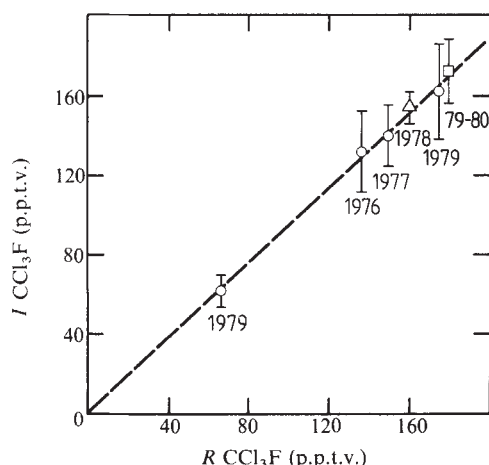


Fig. 2 Mean CCl_3F concentrations, $l(\pm\sigma)$, in air samples analysed by various laboratories, in the years indicated, using independently derived CCl_3F calibration scales, compared with concentrations, R , assigned using the Rasmussen calibration scale (see text). The number of independent laboratories involved in each experiment were 1976, 11; 1977, 16; 1978, 19; 1979–80, 1. A linear best fit through the origin gives $l = 0.952 (\pm 0.040)R$. Data sources: $\circ$, refs 20–22; $\triangle$, ref 23; $\square$, ref 24.

and calibration, the minimum τ_c consistent with observations is 40 yr. The equilibrium atmospheric lifetime (τ_e = equilibrium atmospheric content/equilibrium destruction rate), assuming current releases continue indefinitely, is 66 yr. This agrees well with the most recent two-dimensional model estimates of τ_e (60–65 yr)^{14,15} which also assume that stratospheric photolysis is the only CCl_3F sink. We have shown that this lifetime is, to a first approximation, spatially and temporally consistent with current CCl_3F observations.

The observed and simulated CCl_3F observations for Cape Grim, Tasmania, and Mawson, Antarctica, are shown in Fig. 4. Again the differences between model and observations are a minimum for a τ_c value of 75 yr, the largest discrepancies occurring at the beginning of the observations. The observed temporal trends and model simulations for all sites are listed in Table 1. The South Pole trend ($14.1 \text{ p.p.t.v. yr}^{-1}$) is strongly influenced by the low concentrations measured in January 1975⁵, $\sim 10 \text{ p.p.t.v.}$ lower than the model simulation. Ignoring the initial South Pole data, the globally averaged rate of increase of CCl_3F is $11.9 (\pm 0.7) \text{ p.p.t.v. yr}^{-1}$, in reasonable agreement with the model for the lower troposphere ($11.2 \pm 0.3 \text{ p.p.t.v. yr}^{-1}$).

The rate of change of global CCl_3F concentrations is slowing down. All data sets show d^2C_{obs}/dt^2 (C_{obs} = observed CCl_3F concentration) to be negative (Table 1). Cape Grim data from October 1976 to December 1980 can be represented as a

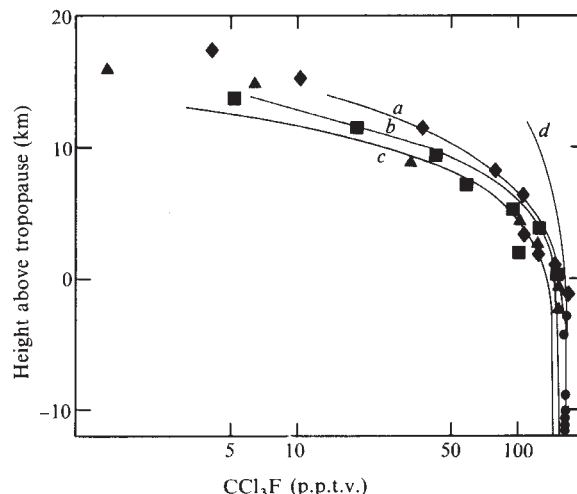


Fig. 3 Observed and calculated vertical distributions of CCl_3F : $\diamond$, 41°N , February 1979; $\square$, 5°S , March 1979; $\triangle$, $78^\circ\text{--}90^\circ \text{S}$, January 1979²⁵; $\bullet$, $30\text{--}40^\circ \text{S}$, September, 1979 (P.J.F., unpublished work). Curves a , b and c are model calculations (mid-latitude Southern Hemisphere 1979) for τ_c values of 100, 75 and 50 yr respectively; d has no CCl_3F stratospheric sink⁵.

quadratic function

$$C_{\text{obs}} = 117.8 + 16.15t - 0.876t^2 \quad (1)$$

where t is in yr, 0 = December 1976. d^2C_{obs}/dt^2 at Cape Grim is $-1.8 \text{ p.p.t.v. yr}^{-2}$, while the globally averaged observed value (Table 1) is $-1.5 (\pm 0.3) \text{ p.p.t.v. yr}^{-2}$. Khalil and Rasmussen¹⁶ observed a rapid decrease in the rate of global increase ($d^2C_{\text{obs}}/dt^2 \sim -3 \text{ p.p.t.v. yr}^{-2}$), based on data from two stations (South Pole and Pacific north-west USA), including the initial South Pole data. Negative values of d^2C_{obs}/dt^2 reflect the reduction in emissions since 1974⁶; however, the model underestimates d^2C_{obs}/dt^2 by a factor of 2. Cape Grim model concentrations from October 1976 to December 1980 can be represented as

$$C_{\text{mod}} = 123.9 + 12.47t - 0.347t^2 \quad (2)$$

d^2C_{mod}/dt^2 at Cape Grim is $-0.7 \text{ p.p.t.v. yr}^{-2}$, and the globally averaged value is $-0.8 (\pm 0.2) \text{ p.p.t.v. yr}^{-2}$ (Table 1). Not only does the model underestimate d^2C_{obs}/dt^2 , but it also over- and under-estimates respectively the initial (late 1976) values of C_{obs} and dC_{obs}/dt (Table 1).

To assess the significance of these discrepancies, sensitivity studies were performed to determine the effect that changes in model parameters have on predictions. The initial qualitative

Table 1 CCl_3F time trends (dC_{obs}/dt), and their rate of change with time from observational data and model simulations ($\tau_c = 75 \text{ yr}$) for various monitoring locations

Station	Period	dC_{obs}/dt (p.p.t.v. yr ⁻¹)	dC_{mod}/dt (p.p.t.v. yr ⁻¹) _n	d^2C_{obs}/dt^2 (p.p.t.v. yr ⁻²)	d^2C_{mod}/dt^2 (p.p.t.v. yr ⁻²)
Point Barrow, Alaska	March 1977–December 1979	11.6(0.5)	10.9(0.4)	-1.8	-0.9
Pacific NW, USA	January 1975–January 1980	11.7(0.6)	11.4(0.3)	-0.9	-0.9
Niwot Ridge, Colorado	March 1977–December 1979	10.5(0.5)	11.1(0.4)	-1.4	-0.6
Mauna Loa, Hawaii	April 1977–December 1979	12.5(0.4)	11.0(0.1)	-1.7	-1.0
Cape Matatula, Samoa	March 1977–December 1979	12.1(0.4)	11.1(0.3)	-1.5	-1.1
Cape Grim, Tasmania	April 1976–December 1980	12.7(0.5)	11.2(0.1)	-1.8	-0.7
	October 1976–December 1980	12.5(0.3)	11.1(0.1)	-1.8	-0.7
Mawson, Antarctica	December 1977–December 1980	11.7(0.7)	11.0(0.2)	-1.2	-0.9
South Pole	January 1975–January 1980	14.0(1.0)	12.0(0.1)	-2.5	-0.5
	January 1976–January 1980	12.6(0.8)	11.9(0.1)	-1.8	-0.6
Global average		11.9(0.7)	11.2(0.3)	-1.5(0.3)	-0.8(0.2)

dC_{obs}/dt and d^2C_{obs}/dt^2 are determined by linear and quadratic regressions respectively. All observational data are from flask air samples, except for Cape Grim, October 1976 to December 1980 (*in situ*, gas chromatography). Global average computed from South Pole data from January 1976, Cape Grim *in situ* data and data from the other six sites.

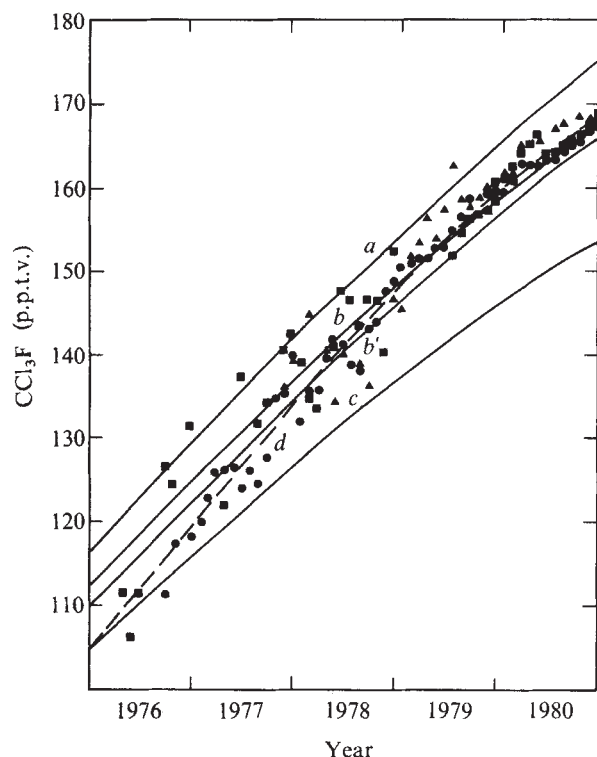


Fig. 4 Observed and calculated monthly mean CCl_3F concentrations at Cape Grim, Tasmania (41°S) (Δ , *in situ* gas chromatograph, 70 samples per month, average monthly standard deviation 2.8 p.p.t.v., $\bullet$, whole air samples analyses at CSIRO, Aspendale, glass flasks, 8 samples per month, average monthly standard deviation 3.4 p.p.t.v.) and Mawson, Antarctica (68°S) ($\blacksquare$, glass flasks, three samples per month, average monthly standard deviation 1.2 p.p.t.v.). Curves *a*, *b* and *c* are model calculations at Cape Grim for τ_c values of 100, 75 and 50 yr respectively. Curve *b'* is for Mawson with $\tau_c = 75$ yr. Curve *d* is the model calculation for Cape Grim with $\tau_c = 75$ yr, assuming a reconstructed release designed to match observations (see text).

studies used a model¹⁷, formulated as

$$dC_{\text{mod}}/dt = -kC_{\text{mod}} + AS(t) \quad (3)$$

where k is the inverse atmospheric residence time, $S(t)$ is a source function and A relates CCl_3F release to concentration. With this model it is clear that the above discrepancies cannot be resolved by changes in the lifetime k^{-1} . In terms of the normalized gradient $\chi_{\text{mod}} = (1/C_{\text{mod}})(dC_{\text{mod}}/dt)$, the model values are smaller than observed. χ_{mod} is independent of k if $S(t)$ increases exponentially¹⁷; for the nearly exponential increase pre-1975, χ_{mod} is so insensitive to changes in k that no positive value can resolve these discrepancies.

To assess the effect release uncertainties have on reducing these discrepancies a constrained inversion formalism¹⁸ was used simultaneously to minimize changes in release data⁶ and discrepancies between the model and observations. Details of the procedure will be published elsewhere. Briefly, to the published release data, $S_0(t)$, are added (or subtracted) a series of 4-yr wide pulses of CCl_3F release $S_n(t)$, centred on 1970, 1971 to 1980, to obtain a modified release,

$$S'(t) = S_0(t) - \sum_{n=0}^{10} a_n S_n(t) \quad (4)$$

The magnitudes (a_n) of these pulses are constrained such that $S'(t)$ is a minimum deviation from $S_0(t)$, that is, $\sum_{n=0}^{10} a_n^2$ is minimized, and are also constrained so that C'_{mod} obtained from the model using $S'(t)$, is a minimum deviation from C_{obs} .

As is normal in inversion problems, many solutions can be obtained depending on the relative weighting attached to minimizing deviations from $S_0(t)$ and C_{obs} . Figure 5 shows a reconstructed release with the weighting chosen to make effective use of both constraints. The cumulative reconstructed

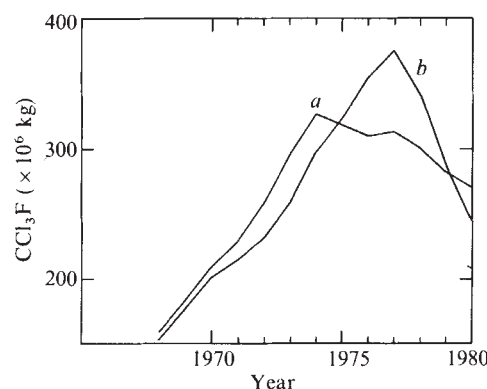


Fig. 5 Global CCl_3F release data (*a*) published⁶ and (*b*) reconstructed as a minimum deviation from (*a*) designed to simulate observations at Cape Grim (see text).

release to 1980 is $4,325 \times 10^6$ kg, 1% lower than published CMA data⁶. A reconstructed release based on data from all observational sites, constrained as above, yields similar results. The differences between $S'(t)$ and $S_0(t)$ are, in general, less than the stated¹⁹ uncertainties in $S_0(t)$ ($\pm 10\%$). However, for a few years $S'(t)$ lies outside these uncertainties, for example, in 1977 $S'(t)$ is 19% higher than $S_0(t)$. The major difference (Fig. 5) is that the maximum emissions occur in 1977 for $S'(t)$ compared with 1974 for $S_0(t)$. The time of the maximum in $S_0(t)$ is determined by when maximum releases of CCl_3F from aerosol products occurred, since releases from other sources grew monotonically through the 1970s. If the currently assumed CCl_3F usage distribution is correct then release $S'(t)$ requires that maximum global aerosol emissions occurred around 1977 rather than 1974.

The model concentrations obtained using the reconstructed release are, as expected, in excellent agreement with observations (Fig. 4) and can be represented by

$$C_{\text{mod}} = 118.0 + 15.73t - 0.778t^2 \quad (5)$$

which agrees well with equation (1). By making a few significant changes to release data C_{obs} , dC_{obs}/dt and d^2C_{obs}/dt^2 have all been fitted simultaneously without any change in tracer transport or destruction. The discrepancies between C_{obs} and C_{mod} cannot be even partially explained with very large changes in the τ_c . It follows that the trend method of estimating atmospheric lifetimes, which attempts to fit the single quantity $1/C_{\text{obs}}(dC_{\text{obs}}/dt)$, will be quite sensitive to errors in the release data.

The gross features of the spatial CCl_3F distribution can be adequately described by a two-dimensional transport model that incorporates published CCl_3F emissions, a realistic photolytic sink and a transport scheme based largely on advection with a small, tuned diffusive component. This model yields τ_c and τ_e values of 75 and 66 yr respectively; the minimum τ_c consistent with observations is 40 yr. Sensitivity studies indicate that more precise determinations of τ_c await a reduction of release uncertainties and the availability of longer, more precise observational records. The detailed temporal changes recorded at Cape Grim cannot be explained within the stated uncertainties of release or by changing the τ_c , always assuming that the discrepancies are real and not due to any time dependent calibration error. These discrepancies can be accounted for if CCl_3F releases from aerosol products reached a higher maximum a few years later than in the published data⁶.

In this reconstruction we have confined our attention to Cape Grim data, as it is the longest, most detailed, continuous baseline CCl_3F record in existence. We intend in the future to assess the sensitivity of derived release rates, based on all available observational data, to variations in interhemispheric transport rates, the relative weights given to the constraints discussed above, and to noise in the observational data.

We thank the Australian Government's Department of Science and Technology and the Department of Meteorology,

University of Melbourne, for observational facilities and technical assistance at Cape Grim, Tasmania, and Mawson, Antarctica, respectively. We acknowledge the advice and assistance of Professor R. A. Rasmussen and Mr A. J. Crawford (Oregon Graduate Center), Professor J. E. Lovelock, University of Reading, Dr J. Peterson (GMCC), and financial assistance of the Aerosol Association of Australia.

Received 17 September 1982; accepted 4 February 1983.

1. Mendonca, B. G. (ed.) *Geophysical Monitoring for Climatic Change*, No. 7 Summary Rep. (1979).
2. Herbert, G. A. (ed.) *Geophysical Monitoring for Climatic Change*, No. 8 Summary Rep. (1980).
3. Rasmussen, R. A., Khalil, M. A. K. & Dalluge, R. A. *Science* **211**, 285–287 (1981).
4. Fraser, P. J. & Pearman, G. I. *Atmos. Envir.* **12**, 839–844 (1978).
5. Hyson, P., Fraser, P. J. & Pearman, G. I. *J. geophys. Res.* **85**, C8, 4443–4456 (1980).
6. *World Production and Release of Chlorofluorocarbons 11 and 12 through 1980*, Revision 1 (Chemical Manufacturers Association, Fluorocarbon Program Panel, February 1982).
7. Plumb, R. A. *J. atmos. Sci.* **36**, 1699–1704 (1979).
8. Pyle, J. A. & Rodgers, C. F. Q. *J. R. met. Soc.* **106**, 421–446 (1980).
9. Schlesinger, M. E. & Mintz, Y. *J. atmos. Sci.* **35**, 1325–1361 (1979).
10. Rowland, F. S. & Molina, M. J. *Rev. Geophys. Space Phys.* **13**, 1–35 (1975).
11. Manabe, S. & Möller, F. *Mon. Weath. Rev.* **89**, 503–532 (1961).
12. WMO Global Zone Research and Monetary Project, Rep. No. 11, *The Stratosphere 1981, Theory and Measurements* (WMO/NASA/FAA/NOAA, 1982).
13. Frederick, J. E. & Mentall, J. E. *Geophys. Res. Lett.* **9**, 461–464 (1982).
14. Ko, M. K. W. & Sze, N. D. *Nature* **297**, 317–319 (1982).
15. Owens, A. J., Steed, J. M., Miller, C., Filkin, D. L. & Jesson, J. P. *Geophys. Res. Lett.* **9**, 700–703 (1982).
16. Khalil, M. A. K. & Rasmussen, R. A. *J. Air Pollut. Control Ass.* **31**, 1274–1275 (1981).
17. Cunnold, P., Alyea, F. & Prinn, R. J. *geophys. Res.* **83**, 5493–5500 (1978).
18. Twomey, S. *Introduction to the Mathematics of Inversion in Remote Sensing and Indirect Measurements* (Elsevier, Amsterdam, 1977).
19. McCarthy, R. L., Bower, F. A. & Jesson, J. P. *Atmos. Envir.* **11**, 491–497 (1977).
20. Rasmussen, R. A., Pierotti, D. J. & Krasnec, J. *Proc. 69th A. Meet. of the Air Pollution Control Association*, Portland, Oregon (1976).
21. Rasmussen, R. A. *Atmos. Envir.* **12**, 2505–2508 (1978).
22. Rasmussen, R. A. & Khalil, M. A. K. *Atmos. Envir.* **15**, 1559–1568 (1981).
23. Fraser, P. J. *WMO Spec. Envir. Rep.* No. 14, 57–63 (WMO No. 549) (1980).
24. Rasmussen, R. A. & Lovelock, J. E. *J. geophys. Res.* (in press).
25. Goldan, P. D., Kuster, W. C., Albritton, D. L. & Schmeltekopf, A. L. *J. geophys. Res.* **85**, 413–423 (1980).

Cooling liquid ^3He to around 100 μK

A. M. Guénault, V. Keith, C. J. Kennedy,
I. E. Miller & G. R. Pickett

Department of Physics, University of Lancaster,
Lancaster LA1 4YB, UK

We have used our immersed-refrigerant cooling method^{1,2}, previously used only to cool ^3He - ^4He solutions^{3,4}, to cool a sample of moderately pure ^3He . In the preliminary experiment reported here we have achieved an extremely low helium temperature, certainly below 120 μK , at which temperature the superfluid B phase is so far below its transition temperature that our vibrating wire viscometer is essentially unable to detect the presence of any normal fluid component. In this regime we hope to be able to investigate any interaction between the wire and the stiffness of the superfluid texture.

During the past 18 months we have been cooling dilute solutions of ^3He in ^4He below 0.4 mK by adiabatic demagnetization of nuclear spins in copper^{1–4}. Since the Kapitza coupling between metals and dilute ^3He - ^4He solutions is very much worse than that between metals and pure ^3He several improvements have to be made to cool dilute solutions to these temperatures. The principal lines of attack have been to minimize the stray heat leak into the helium and the thermal path between helium and refrigerant. The former is achieved by using a double cell with the specimen volume of helium in the inner cell suspended inside a larger volume in the outer cell acting as a thermal guard. The refrigerant is put into direct contact with the helium by immersing it in the liquid sample with a very large area of surface contact. The surface is provided either by using finely divided copper particles or by using slabs of copper coated with sintered silver powder.

A major difficulty in these experiments has been that of thermometry, since the same poor Kapitza coupling which militates against cooling the specimen also militates against cooling the thermometer. Our usual thermometric quantity is the nuclear susceptibility of platinum metal measured by pulsed NMR using a commercially available system⁵. We have tried

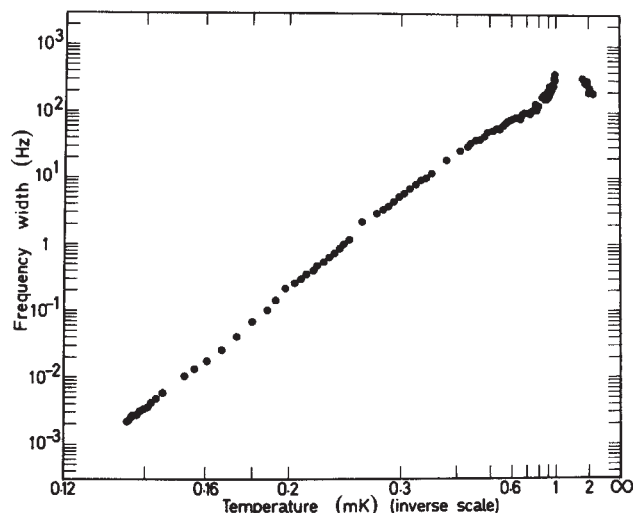


Fig. 1 A series of measurements of the frequency width of the vibrating wire versus temperature in liquid ^3He at a pressure of 0.0 bar. Because of the low magnetic field (0.014 T), measurements of the large Δf_2 values near the transition temperature (1.04 mK) were not practicable; the temperature scale is therefore adjusted to fit a calculation of the known behaviour of Δf_2 in the normal fluid above T_c , but this preliminary scale could be in error by as much as 10%. (An internal calibration of the Pt thermometer against the spin-lattice relaxation time τ_1 was not possible in the present wire thermometer owing to problems arising from poor insulation of the wires.)

several configurations for making thermal contact to the platinum: using finely divided platinum powder, fine powder diluted with an insulator (to reduce eddy current heating) or wires thermally anchored to silver sinter pads immersed in the liquid. All methods have their disadvantages and we have not so far produced a thermometer which will return reliable temperatures below 0.35 mK in the dilute solutions (even though we have evidence that the liquid is colder). A further nontrivial problem is the calibration of such thermometers. In experiments on ^3He - ^4He solutions we have no obvious fixed points for calibration (such as the superfluid transition in pure ^3He) and at higher temperatures where spin-lattice relaxation measurements are reasonably reliable our small thermometers tend to be too insensitive for adequate calibration.

To avoid these problems we have recently tried cooling pure ^3He so as to be able to calibrate the current thermometer using the superfluid transition at low pressure around 1 mK as a fixed point, together with the better known behaviour of the viscosity in normal ^3He above the transition.

On cooling ^3He we find that our system, designed to overcome the difficulties associated with dilute solutions, will readily cool pure ^3He to very low temperatures. During a series of preliminary demagnetizations during the 10 days before Christmas 1982 we cooled the B phase at 0.3 bar to around 120 μK , or to be more precise we have cooled the platinum wire and silver sinter thermometer immersed in the liquid to about 130 μK while the liquid appears to be considerably colder. (The ^3He was used in the as-received state with a ^3He content of $\sim 0.1\%$.)

The inner cell is furnished with a vibrating wire viscometer which consists of a length of 0.124 mm diameter tantalum wire bowed into an approximate semicircle of radius 4 mm contained in a free volume of $\sim 10 \times 10 \times 3$ mm. A steady magnetic field of 14 mT is applied in the plane of the semicircle. The natural flapping resonance of the loop is excited by the Lorentz force on the wire as it carries an alternating current. The amplitude of the resonance of the wire is measured from the voltage induced by the movement in the field. Thus the frequency shift and the frequency width can be observed by the slow sweeping of the frequency of the driving current through the resonance. The frequency shift primarily indicates how much liquid moves with the wire, whilst the width of the resonance is a measure of the damping.

Cooling the ^3He to the very low temperatures has a spectacular effect on the frequency width, Δf_2 , of the resonance. At the superfluid transition ($T_c \approx 1.1$ mK at 0.3 bar) Δf_2 is estimated to be around 400 Hz for our wire, which has a vacuum frequency for 890 Hz. At the lowest temperatures reached the width has fallen below 0.004 Hz. Since virtually all of the observed width is accounted for by the vacuum damping of the wire alone, we estimate that the damping from the liquid falls to a value below 10^{-3} Hz, representing a fall in damping of more than five orders of magnitude from the value at T_c . This behaviour is illustrated in Fig. 1, which plots measurements of Δf_2 against temperature.

The measurements at the lowest temperatures are not straightforward to make. The driving level of the vibrating wire must be kept below the value at which heating effects can be observed, and the sweep time must be very long (at least 2,500 s) since the ringing time of the resonator is several minutes.

The data in Fig. 1 have been corrected for an estimate of the vacuum damping of the wire, $\Delta f_2^{(0)} = 0.003$ Hz. Paradoxically, this quantity is not easy to measure independently, because the excitation gas at 130 μK in ^3He represents a much better 'vacuum' (the calculated bulk mean free path is measured in tens of metres) than any we can hope to achieve by pumping. Hence we have estimated $\Delta f_2^{(0)}$ from the asymptotic behaviour of Δf_2 -T at low temperatures, although in fact the precise value only affects the data in Fig. 1 at widths below ~ 0.03 Hz.

Below 170 μK the ^3He quasiparticles play a negligible role, and we enter a hitherto unstudied regime in which any mechanical response of the vibrating wire to the helium is determined only by the superfluid condensate. A significant response is possible in this case because the superfluid ordering in liquid ^3He , unlike that in ^4He , has directional properties or 'texture'. We have already observed indications that this superfluid texture does indeed have a mechanical influence on the movement of the wire, in that the wire ceases to behave as a classical harmonic oscillator below 170 μK . Nonlinear effects appear, giving both an asymmetric lineshape to the resonance and also a shift of the observed resonance frequency with driving amplitude. These effects, currently under more detailed study, cannot be ascribed merely to heating as the shift increases faster with decreasing amplitude.

Received 20 February; accepted 8 March 1983.

1. Dow R. C. M., Guénault A. M. & Pickett G. R. *J. Low Temp. Phys.* **47**, 477-489 (1982).
2. Bradley D. I. *et al. Cryogenics* **22**, 296-304 (1982).
3. Bradley D. I., Guénault A. M., Keith V., Pickett G. R. & Pratt W. P. *J. Phys. C: Solid St. Phys.* **30**, L501-L506 (1982).
4. Guénault A. M., Keith V., Kennedy C. J. & Pickett G. R. *Phys. Lett.* **90A**, 432-434 (1982); *Phys. Rev. Lett.* **50**, 522-525 (1983).
5. *Instruments for Technology* (PLM 2).

Rates of sedimentation, and stratigraphical completeness

John C. Tipper

Department of Geology, University College Galway, Galway, Republic of Ireland

The most obvious result of sedimentation processes is the change in the thickness of accumulated sediment with time. Whenever the net change over a period of time is positive, a record is left of a sedimentation system's existence, a stratigraphical section. For the stratigrapher, the problem that such a section poses is to determine the detailed correspondence of thickness to time¹; for the sedimentologist, the problem is to infer the nature of the parent sedimentation system, and to estimate its parameters. A simple stochastic model is used here: (1) to illustrate some basic stratigraphical results of clastic sedimentation processes; (2) to offer an explanation for the commonly cited systematic decrease of mean sedimentation rate with increasing time span; and (3) to assess the concept of stratigraphical completeness. (It will be assumed that all post-depositional, non-erosional processes, such as compaction, can be identified, and their effects on stratigraphical thickness allowed for.)

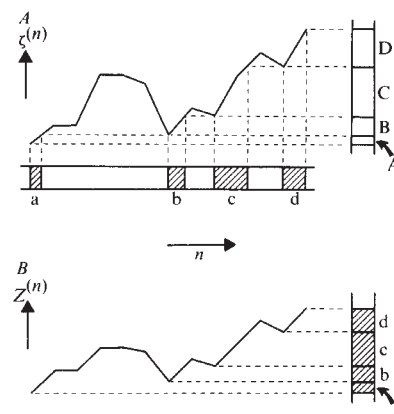


Fig. 1 A, Relationship between thickness, $z^{(n)}$, and time, n , for a sedimentation system of which four beds, A-D, are the stratigraphical record (vertical column). Corresponding represented time intervals, a-d, are shaded in the time section (horizontal). B, Relationship between net duration of deposition, $Z^{(n)}$, and time, n , for the same sedimentation system. Represented intervals, a-d, in the time record (vertical column) are identical to those in the time section of the thickness-time diagram.

Most stratigraphical sections are manifestly discontinuous, at all scales. Sedimentation units have laminae within: formations are commonly subdivisible into members. Stratigraphical homogeneity is relative, discontinuity more or less evident. Each depends on the scale at which it is viewed, yet each reflects the underlying unsteadiness of the sedimentation system that gave rise to it.

Typically, in most clastic sedimentation systems, states of significant deposition and erosion are relatively brief, uncommon episodes—non-deposition, or at least the absence of readily discernible deposition and erosion, is normal. Typically too, when they do occur, the intensities of the depositional and erosional processes (the instantaneous rates of deposition and erosion) vary in time from each episode to its successor. The parameters of the sedimentation system, that is the relative frequency and duration of episodes of deposition and erosion, the pattern of their succession, and the intensity of the depositional and erosional processes, determine the way in which sediment accumulates from that system, if at all (Fig. 1A).

Perhaps the simplest stochastic model that can reasonably describe such systems is the general one-dimensional random walk²⁻⁶. It may be applied by letting the typical episodes of deposition and erosion define minimal units of a discrete time scale: the lengthy periods of non-deposition, as well as any prolonged intervals of deposition or erosion, are treated just as multiples of these units. Within the r th time unit the thickness change is δ_r , and this is distributed according to a frequency function, $f(\delta)$, independent of t (Fig. 2A). Members of the set $\{\delta_r\}$ are thus mutually independent, identically distributed random variables.

The expected value of δ , $E(\delta)$, defines the two regimes of sedimentation: a depositional regime has $E(\delta)$ positive, an erosional regime has $E(\delta)$ negative. A real stratigraphical section is defined by $z^{(n)} > 0$, where $z^{(n)} (= \sum_{i=1}^n \delta_i)$ is the net accumulation over n time units. (Note that the existence of a real stratigraphical section does not imply that the regime of its parent system was depositional: net accumulation is quite common in erosional regimes, and net erosion in depositional ones.)

The conditional probability distribution, $\text{Prob}(z^{(n)}|n)$, describes the thickness-time correspondence for any particular system, that is, for a particular $f(\delta)$. Yet irrespective of the form of $f(\delta)$, $\text{Prob}(z^{(n)}|n)$ is approximately normal for large n , in this context for sections that span time intervals that are long with respect to the typical depositional or erosional episode. If

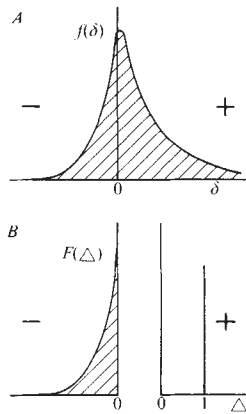


Fig. 2 A, Frequency function, $f(\delta)$, of thickness change, δ . System illustrated is of depositional regime, that is $E(\delta) > 0$. B, Frequency function, $F(\Delta)$, of change in duration of preserved deposition, Δ , for the same sedimentation system. For $\Delta \leq 0$, $F(\Delta)$ is continuous; for $\Delta > 0$, $F(\Delta)$ is a single spike at $\Delta = 1$.
 $\int_{-\infty}^0 f(\delta) d\delta = \int_{-\infty}^0 F(\Delta) d\Delta = \int_0^{\infty} f(\delta) d\delta = \text{Prob}(\Delta = 1)$.

the first and second moments of $f(\delta)$ are μ and σ^2 , then $\text{Prob}(\zeta^{(n)}|n)$ is $N(n\mu, n\sigma^2)^2$. The expected rate of net thickness change over n time units, or what is commonly understood by the mean rate of sedimentation (for depositional regimes) or the mean rate of erosion (for erosional regimes), is then simply $E(\zeta^{(n)}|n)$. The model predicts that this rate, R , is equal to μ and hence independent of n , the time span over which it applies.

This result is important because it conflicts with the overwhelming geological evidence that mean rates of sedimentation estimated from the stratigraphical record decrease systematically with the time span over which they are determined⁷. Thus, if $\zeta_i^{(n)}$ is the measured thickness corresponding to a time interval of length n in the i th of m sections, $\hat{R}_n (= \sum_{i=1}^m \zeta_i^{(n)} / mn)$ is found to decrease with n . But what decreases with n is not the mean rate of sedimentation, for $\hat{R}_n$ is not in any sense an acceptable estimator of R (except for entirely non-erosional systems).

The relationship between R and $\hat{R}_n$ is quite straightforward. $\hat{R}_n$ is really an estimator of $E(\zeta^{(n)}|n, \zeta^{(n)} > 0)$, because $\hat{R}_n$ can only of course be determined for real stratigraphical sections ($\zeta_i^{(n)} > 0$, for all i). From the model, $E(\zeta^{(n)}|n, \zeta^{(n)} > 0)$ is just the expected value of the positive section of $N(n\mu, n\sigma^2)$, truncated at zero. $\hat{R}_n$ is then theoretically equivalent to $R + (\sigma/\sqrt{2\pi n})$ (Fig. 3). The systematically decreasing relationship of $\hat{R}_n$ to n may be due to nothing more than the fact that the stratigrapher can never use other than real stratigraphical sections.

The random walk model is also useful for assessing the concept of stratigraphical completeness^{7,8}. For a section spanning n time units, the completeness, C_n , may be defined in terms of $Z^{(n)}$, the net duration of deposition represented by the section. $Z^{(n)}$ is the sum of random variables Δ_t , $t = 1, \dots, n$ (Fig. 1B), where Δ_t , the change in duration of preserved deposition during the t th time unit, is distributed according to some frequency function, $F(\Delta)$, independent of t (Fig. 2B). C_n , which is non-zero only for real stratigraphical sections, is defined as:

$$\begin{aligned} Z^{(n)}/n & \text{ for } Z^{(n)} > 0 \\ 0 & \text{ for } Z^{(n)} \leq 0 \end{aligned}$$

As for $\{\delta_i\}$, $\{\Delta_i\}$ are mutually independent, identically distributed random variables, and as for $\zeta^{(n)}$, $Z^{(n)}$ is normally distributed for large n , in this case $N(n\nu, n\rho^2)$, where ν and ρ^2 are the first two moments of $F(\Delta)$. The expected stratigraphical completeness for a section spanning n time units, $E(C_n|n)$, is thus $\nu + (\rho/\sqrt{2\pi n})$, and has the same type of systematically decreasing relationship with n as does $\hat{R}_n$ (Fig. 4A).

It is also possible to consider stratigraphical completeness in terms of intervals of time that are greater than the basic time units⁷: this may be termed interval completeness. If m intervals can be accommodated in the time span of a section, each interval of k time units, then the interval completeness of the section, $C_n(k)$, is defined as m^*/m , where m^* is the number of intervals that are to some extent represented in the section, that is for which $Z^{(k)} > 0$. k is effectively a measure of the stratigraphical resolution.

The distribution of m^* is binomial with parameter p_k , the probability that any interval of length k will be represented. Hence $E(C_n(k)|n, k)$ is p_k . For large k (for low stratigraphical resolution) the distribution of $Z^{(k)}$ is $N(k\nu, k\rho^2)$ and p_k thus $\Phi(\nu\sqrt{k}/\rho)$, where $\Phi(x)$ is the standardized normal distribution function. For small k , the normal approximation does not hold and the value of p_k depends on the form of $F(\Delta)$. An upper bound for p_k , however, is $1 - (1 - \alpha)^k$, where α is $\text{Prob}(\Delta = 1)$. (Fig. 4B). In either case, the expected interval completeness of a section is controlled just by the nature of the parent sedimentation system and by the required stratigraphical resolution. In contrast to C_n , it is independent of n , the section's time span.

Because in the random walk model, the two measures of stratigraphical completeness differ so fundamentally in their behaviour, we now question both the practical validity of the model's assumptions, and the real meaning of stratigraphical completeness. There are three assumptions: (1) that deposition and erosion rates can be considered constant within each of the discrete minimal time units (Fig. 1A); (2) that successive episodes of deposition or erosion are mutually independent; and (3) that the parameters of the model are invariant with time, that is the model is stationary. Of these, the definition of the minimal time unit is critical, for within it deposition (or erosion) must be sensibly steady and from any time unit to its successors there must be neither significant dependence nor systematic trend of deposition or erosion rates. How then might such minimal time units be defined for real sedimentation systems and real stratigraphical sections?

For the sedimentologist, the appropriate duration of time unit will inevitably be short, to accord with the observable sedimentation processes. In contrast, the stratigrapher will tend to select a substantially longer time unit to emphasize the independence of successive depositional or erosional events, notwithstanding their innate unsteadiness. By the same token,

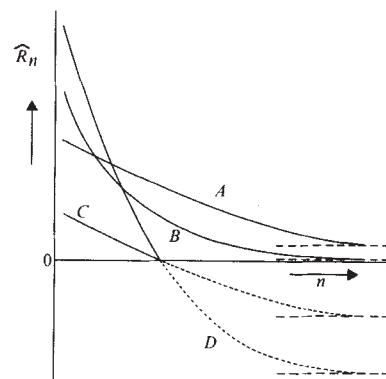


Fig. 3 Relationship between $\hat{R}_n$, the mean rate of sedimentation estimated from real stratigraphical sections of duration n , and n . A, Depositional regime; low variability of rates of deposition/erosion, see, for example, the system of Fig. 2. B, Depositional regime; high variability. C, Erosional regime; low variability. D, Erosional regime; high variability. In the depositional regime the value of R , the expected rate of sedimentation, is the value to which $\hat{R}_n$ asymptotes for large values of n . In the erosional regime this asymptotic relationship can never be demonstrated in practice, because $\hat{R}_n$ can never be negative.

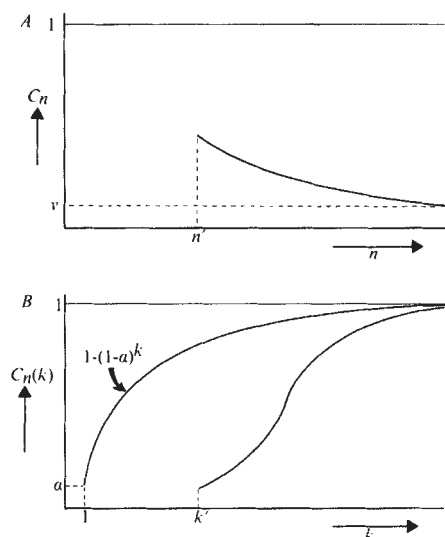


Fig. 4 *A*, Relationship between stratigraphical completeness, C_n , and time duration, n , for sedimentation system of Fig. 2. n' is the time duration above which the normal approximation holds. C_n asymptotes to v , the mean value of Δ (Fig. 2*B*). *B*, Relationship between interval completeness, $C_n(k)$, and interval duration, k , for the same sedimentation system; k' is the interval duration above which the normal approximation holds. Curve of $1 - (1 - \alpha)^k$ is the upper bound for $C_n(k)$, where $\alpha = \text{Prob}(\Delta = 1)$ (Fig. 2*B*).

the sedimentologist will take a more restrictive view of the stationarity assumption than will the stratigrapher. (Note the further assumption that it is useful to treat any particular sedimentation system as 'lasting' for any length of time. Ultimately every sedimentation system is specific both to its time and to its location; it has but one, unique realization. Concepts such as mean rate of sedimentation can refer at best only to a family of like systems.)

With the minimal time unit selected, by whatever criteria, the model can be used to predict stratigraphical completeness, as measured either by C_n or $C_n(k)$. When C_n is used, the stratigraphical resolution is that of the minimal time unit; for $C_n(k)$ it is that of the interval length k . Comparable resolution is evidently attainable by calculation $C_n(k)$ for a shorter minimal time unit (of duration, say, t), and by using C_n based on a longer minimal time unit (of duration kt). Yet the results conflict.

Of the two measures of stratigraphical completeness, C_n is the more realistic, because it alone is dependent on the time span of the stratigraphical section. To calculate its value for a given n , the form and parametrization of $f(\delta)$, and hence of $F(\Delta)$, must first be determined. The moments of $f(\delta)$, $\mu (=R)$ and σ^2 , can be estimated from values of $\bar{R}$, but its form can only be found by analogy with comparable modern sedimentation systems. As is generally the case, it is only by analysis and reconstruction of the parent sedimentation system that its stratigraphical results can be appreciated.

I thank Dr I. Ó. Muirchearthaigh for advice and assistance. This paper is a contribution to IGCP project 175.

Received 19 November 1982; accepted 25 February 1983.

Noble gas constraints on the layered structure of the mantle

Ichiro Kaneoka

Geophysical Institute, Faculty of Science, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan

Several chemical models of mantle structure have been proposed¹⁻⁶ based on isotopic data such as Sr, Nd, Hf and Pb in volcanic rocks from both oceanic and continental areas. The most popular of these are two chemical models of the vertically-layered mantle: (1) that it is a depleted mantle from which mid-oceanic ridge basalt (MORB) is derived and (2) that it is a fertile mantle from which a mantle plume may arise¹⁻⁵. However, these models are almost always based on solid elements, such as Sr, Nd, Hf and Pb isotopes, and do not fully reflect the noble gas data. Because noble gases are liable to be lost from the solid earth when a mantle differentiation occurs on a global scale in the upper mantle, the data on noble gas isotopes could give constraints on the structure of the mantle that differ from those by solid element isotopes. Based on a consideration ³He/⁴He and ⁴⁰Ar/³⁶Ar ratios of recent samples, I argue here that the observed data suggest that mixing occurs between the source materials of a mantle plume and of MORB. The data seem to favour the model which assumes that the source of a typical mantle plume is deeper than that of MORB.

It is not necessary to assume a layered mantle when explaining the occurrence of isotopically different sources in the mantle. In effect, a mantle model is proposed where fertile parts are distributed like spots in the depleted mantle⁶. However, this model is valid only when the isotopic disequilibrium is maintained on rather local scale of, say, <100 km for >1,000 Myr. As shown below, even the noble gas data, including He, require the existence of the fertile mantle together with the depleted mantle. At present, we have no appropriate data on the diffusion of noble gases for mantle materials. However, Jambon and Shelby reported high diffusion coefficients of $\sim 10^{-3} \text{ cm}^2 \text{ s}^{-1}$ even at 300 °C for He in obsidians⁷. If we assume that the mantle materials have the same order of diffusion coefficients for He in the mantle, it means the movement of He for >50 km during 1,000 Myr occurs by diffusion only. Furthermore, if the speed of a convection cell in the mantle is approximated to that of the present plate motion, the total movement of one spot in a convection cell becomes of the order of 10^4 – 10^5 km for 1,000 Myr. Hence as long as we consider the coupling effects of both the mass transport by convection and the relatively large mobility of noble gases such as He, it is difficult to imagine that a locally heterogeneous mantle concerning He and some other volatile elements could be kept as a closed system for more than several hundred million years. These conjectures require the presence at least of the two mantle sources which are chemically different on global scale and might form the layered mantle. This does not always imply the occurrence of homogeneity in the mantle for solid elements on local scale of the order of 100 km. To discuss the problem of magma genesis, we should also consider the detailed structure of the mantle. However, the main concern here is to identify the chemical structure of the mantle on a global scale which might have existed for more than 1,000–2,000 Myr and not the local heterogeneity in the mantle which might have been produced through some magmatic processes and maintained for some limited period of <1,000 Myr, for example. The minimum requisite for a mantle model based on both the solid-element and noble-gas isotopes is the occurrence of the two mantle sources which are chemically different on a global scale. Knowledge of the geometrical relationship between the two mantle sources is one of the most important bases on which to construct a mantle evolution model. Since light noble gases are more mobile than solid elements, they would be more easily

1. Schwarzscher, W. *Random Processes in Geology* (ed. Merriam, D. F.) 96–111 (Springer, New York, 1976).
2. Cox, D. R. & Miller, H. D. *The Theory of Stochastic Processes* (Chapman & Hall, London, 1965).
3. Kolmogorov, A. N. *Am. math. Soc. Transl.* No. 53 (1951).
4. Mizutani, S. & Hattori, I. *Math. Geol.* **4**, 123–146 (1972).
5. Schwarzscher, W. *Sedimentation Models and Quantitative Stratigraphy* (Elsevier, Amsterdam, 1975).
6. Dacey, M. F. *Math. Geol.* **11**, 655–668 (1979).
7. Sadler, P. M. *J. Geol.* **89**, 569–584 (1981).
8. Van Andel, Tj. H. *Nature* **294**, 397–398 (1981).

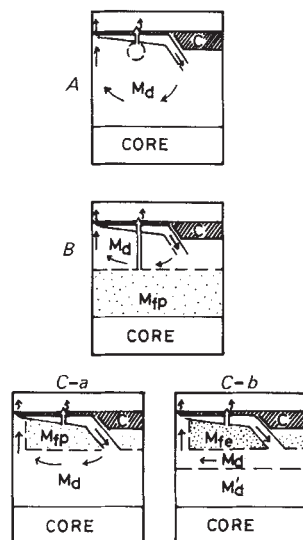


Fig. 1 Models of the chemically layered structure of the mantle. Because the isotopic data require at least more than one source materials in the mantle, model A is not supported and not discussed in detail in the text. M_d , depleted mantle; M_{fp} , fertile mantle (primitive); M_{fe} , fertile mantle (enriched); C, crust.

homogenized than the latter and the noble-gas isotope system would be more appropriate for inferring the mantle structure on global scale. Hence, here we examine the two layered-mantle models from the viewpoint of noble-gas isotopes and disregard any mantle evolution model which requires more assumptions than can be derived from the present data.

There are two essentially different layered mantle models. The first assumes that the upper mantle is depleted, while the lower mantle is fertile¹. The fertile mantle and the depleted mantle are classified here only in a relative sense that the incompatible elements such as K and U are more abundant in the former than the latter. The second model assumes that the upper mantle is fertile and the lower mantle is depleted. This model can be further divided into two alternatives: that the fertile mantle is the remains of a primitive one³, or that the enriched one has been caused by mantle differentiation⁴. These models are shown in Fig. 1 where a mantle model with only the depleted mantle is included for comparison.

The noble gas isotopes, $^3\text{He}/^4\text{He}$ and $^{40}\text{Ar}/^{36}\text{Ar}$ ratios are the most useful for identifying different source materials, because they include time-integrated radiogenic components such as ^4He and ^{40}Ar . Kaneoka and Takaoka⁸ concluded that the source material of Hawaiian volcanic rocks should be different from that of MORB and that mixing of two sources is likely, on the basis of the observation that phenocrysts in volcanic rocks show systematically higher $^3\text{He}/^4\text{He}$ and lower $^{40}\text{Ar}/^{36}\text{Ar}$ ratios than those of MORB. Large olivine and pyroxene phenocrysts can trap magmatic noble gases even if groundmass is equilibrated with the atmospheric noble gases at the time of eruption, because such large phenocrysts are thought to be formed in a magma reservoir. Subsequent studies have confirmed the usefulness of such phenocrysts for investigating the magmatic noble gases for samples from hotspot areas⁹.

In Fig. 2, $^3\text{He}/^4\text{He}$ and $^{40}\text{Ar}/^{36}\text{Ar}$ ratios for various samples of recent ages are summarized, and olivine and pyroxene phenocrysts of volcanic rocks from hotspot areas such as Hawaii, Iceland and Réunion, pillow basalts for MORB, volcanic and hydrothermal gases are included. These data are most easily explained by mixing of at least four types of sources of noble gases: MORB(M), plume(P), atmosphere(A) and crust(C). Of these P-type and M-type are directly related to the mantle source and only these will be discussed here.

P-type is characterized by high $^3\text{He}/^4\text{He}$ and low $^{40}\text{Ar}/^{36}\text{Ar}$ ratios as observed in samples from hotspot areas. M-type ratios

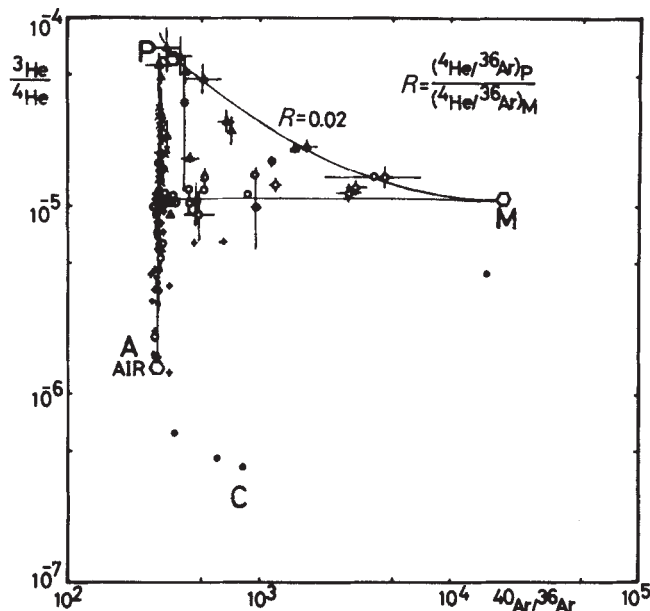


Fig. 2 The $^3\text{He}/^4\text{He}$ – $^{40}\text{Ar}/^{36}\text{Ar}$ plot for samples of recent ages. Solid lines indicate the calculated mixing lines among different typical source materials. P, Plume-type; M, MORB-type; A, atmosphere-type; C, crust-type. Data were compiled from refs 8, 9, 13–15. $\circ$, MORB; $\bullet$, pillow basalt (Kilauea rift zone); $\blacktriangle$, lava, phenocrysts (plume area). Gas: $\times$, Iceland; $+$, arc; $\bullet$, continent.

are represented by those of MORB. In Fig. 2, three mixing lines are calculated, which pass through P, M and A. P, M and A are assumed to represent typical values for the P-type source material ($^3\text{He}/^4\text{He} = 6 \times 10^{-5}$, $^{40}\text{Ar}/^{36}\text{Ar} = 350$), the M-type source material ($^3\text{He}/^4\text{He} = 1.1 \times 10^{-5}$, $^{40}\text{Ar}/^{36}\text{Ar} = 20,000$) and the atmosphere ($^3\text{He}/^4\text{He} = 1.4 \times 10^{-6}$, $^{40}\text{Ar}/^{36}\text{Ar} = 295.5$), respectively. Although the values assigned for the P-type and M-type source materials are only approximations for the true values uncontaminated by the other component, it does not affect the following discussion.

In Fig. 2, most data are located in the area surrounded by the three mixing lines. Those which show lower $^3\text{He}/^4\text{He}$ ratios of $<1 \times 10^{-5}$ together with higher $^{40}\text{Ar}/^{36}\text{Ar}$ ratios than the atmospheric value probably reflect the effect of crustal materials. When a relatively low $^{40}\text{Ar}/^{36}\text{Ar}$ ratio close to the value of the atmospheric Ar is observed, the low value is often attributed to the atmospheric contamination. In hotspot area samples, including those from Hawaii and Iceland, however, higher $^3\text{He}/^4\text{He}$ ratios than that of MORB are always accompanied by relatively lower $^{40}\text{Ar}/^{36}\text{Ar}$ ratios. Furthermore, some samples are located around the mixing line between P and M, suggesting a mixture of source materials. In Fig. 2, R is defined as $(^3\text{He}/^4\text{He})_P / (^3\text{He}/^4\text{He})_M$. For the mixing line between P and M, $R = 0.02 (\pm 0.005)$ is the best fit. By combining the values of P and M with the values for R , we can obtain a value for $(^3\text{He}/^4\text{He})_P / (^3\text{He}/^4\text{He})_M$ of $1.1 (\pm 0.03)$. R can approximate that of the source materials as long as similar magmatic processes control both the P-type and M-type source materials.

Based on the results shown in Fig. 2, we can set a limit for the relative abundance of primordial components ^3He and ^{36}Ar in the source materials for P and M. By assuming the points P and M and a mixing line which go through P and M, we can define a relationship between the $(^3\text{He})_P / (^3\text{He})_M$ and the $(^{36}\text{Ar})_P / (^{36}\text{Ar})_M$ ratios as $(^3\text{He})_P / (^3\text{He})_M = k (^{36}\text{Ar})_P / (^{36}\text{Ar})_M$ where $k = (^3\text{He}/^4\text{He})_P / (^3\text{He}/^4\text{He})_M R$. By inserting appropriate values, we obtain $k = 0.11 (\pm 0.03)$. As the M-type source material is depleted in incompatible elements including U and K compared with that of the P-type source material, the concentrations of radiogenic components such as ^4He and ^{40}Ar in the P-type source material should be larger than those in the

M-type source material, unless the retention age of these components is much longer for M than that for P. Hence, by assuming that P and M in Fig. 2 represent the values for the source materials for the P-type and M-type, we can infer that the minimum values for the $(^3\text{He})_P/(^3\text{He})_M$ and the $(^{36}\text{Ar})_P/(^{36}\text{Ar})_M$ ratios are 5.5 and 57, respectively. This indicates that the P-type source material should contain larger amounts of primordial noble gas components compared with that of M-type. The apparent difference in the degree of retention between He and Ar in each source probably reflects their different mobilities.

To calculate the mixing rate between the two source materials, we need to know the isotopic ratios and relative concentrations for each endmember. Although P and M do not necessarily represent the endmembers, we assume that P and M approximate the values for each source material and that the relative concentrations of noble gases are represented by the minimum values stated above. In this case, we can get a result which corresponds to the minimum mixing rate for the M-type source material with the P-type source material. Even here, however, if a sample shows a $^3\text{He}/^4\text{He}$ ratio of 2×10^{-5} as in the case of Kilauea samples¹⁰, ~70% of the M-type source material might have mixed with 30% of the P-type source material. In general, the mixing rate of the M-type should be larger than this estimate, which means that Kilauea samples might have been produced by mixing more than double the amount of noble gas in M-type source material with that of P-type source material. These inferences give us constraints on the layered structure model of the mantle.

First, if the fertile mantle is located above the depleted mantle, such as in models C-a and C-b, an oceanic plate over it is considered to be the only source for the M-type source material. As long as we allow the mixing process to explain the observed data in Fig. 2, a mantle plume derived from the P-type source should melt more than double the M-type source material in the case for Kilauea samples. However, from the viewpoint of energy balance it is unlikely that a mantle plume could melt more than double the amount of lithosphere material.

Second, we observe large amounts of primordial noble gases in the fertile mantle. In models B and C-a, these are regarded as representing the primitive mantles, whereas in model C-b they are considered to be the result of an enrichment process such as mantle differentiation. In model C-b, the crust and the enriched mantle were formed through mantle differentiation producing the depleted mantle⁴. In this case, there is no clear distinction between the crust and the enriched mantle. To explain the difference in the observed $^3\text{He}/^4\text{He}$ ratios between the P-type and M-type sources by the enrichment process, a simple calculation indicates that such enrichment should have occurred before more than 2,000–2,500 Myr ago. As shown in Fig. 2, the continental crust seems to have lost most of its primordial noble gases. The high $^{40}\text{Ar}/^{36}\text{Ar}$ ratio for the M-type source also requires more than 60–70% of degassing of primordial components before more than 2,000 Myr ago from the mantle of M-type. Such degassing might have occurred through the combination of large-scale mass transport and the relatively large mobility of noble gases. Although we do not yet have sufficient information on the detailed mechanism of degassing from the Earth's interior, we consider that the mantle differentiation on global scale is probably accompanied by extensive degassing as long as the depleted mantle formed secondarily. Furthermore, if the P-type source was produced by the enrichment process and is located in the upper mantle without degassing, we cannot explain the separation of the P-type and the C-type materials as classified in Fig. 2. This implies the occurrence of rather different systems for these components. Hence, the enrichment process is not appropriate for explaining the abundant primordial noble gas contents in the P-type source material.

Third, mantle convection is likely to have caused large-scale mass transport. If it started early in the Earth's history and had

a speed similar to that of the present plate movement, more than 10 rotations of cells would be expected to have occurred in the layered mantle. Hence, taking into account both the mass transport by mantle convection and the relatively large mobility of light noble gases such as He, we consider it difficult for light noble gases to have been retained in a limited region of <100 km for more than several hundred million years. Helium at least would be degassed to the atmosphere unless it were retained at relatively deeper level such as the lower mantle.

Fourth, the $^3\text{He}/^4\text{He}$ ratios for MORB are reported to be relatively uniform over the Pacific, Atlantic and Indian Oceans¹¹. To produce such a uniformity in the $^3\text{He}/^4\text{He}$ ratio, a uniform $^3\text{He}/(\text{U}+\text{Th})$ ratio or some homogenized process such as a low-velocity zone⁸ is required in the mantle. He, and probably Ar, would be more easily homogenized through a partially melting part than through the solid part. On the other hand, samples of the P-type show more variation although they have systematically higher $^3\text{He}/^4\text{He}$ ratios than the MORB values. The variation can be most easily explained by model B, in which a mantle plume arises through the depleted mantle of M-type and it would be affected by mixing with the M-type source material in different degrees. Because the M-type material can extrude with no effect except for the surface materials, it will show much more uniformity in the isotopic ratios.

These conjectures lead us to conclude that noble gas isotope data are most compatible with model B out of the four shown in Fig. 1. In this case, a mantle plume at typical hotspot areas such as Hawaii and Iceland should arise from a relatively deeper mantle, though its depth cannot be determined from the present results. On the other hand, recent studies have revealed another type of hotspot such as Gough and Tristan da Cunha, where the samples show lower $^3\text{He}/^4\text{He}$ ratios than do the MORB values¹². These source materials would probably have been affected by a secondary process, such as contamination by crustal materials. Hence, only a typical mantle plume whose source is represented by a higher $^3\text{He}/^4\text{He}$ ratio compared with that of MORB probably arises from the relatively deeper part of the mantle which still retains primordial noble gases.

I thank Professor O. L. Anderson for valuable comments on the original manuscript. This study is financially supported by the Ministry of Education, Science and Culture under the contract 57540204.

Received 16 August 1982; accepted 17 February 1983.

- Jacobsen, S. B. & Wasserburg, G. J. *J. geophys. Res.* **84**, 7411 (1979).
- O'Nions, R. K., Evensen, N. M. & Hamilton, P. J. *J. geophys. Res.* **84**, 6091 (1979).
- Tatsumoto, M. *Earth planet. Sci. Lett.* **38**, 63 (1978).
- Anderson, O. L. *Earth planet. Sci. Lett.* **57**, 13 (1982).
- Allègre, C. J., Brévart, O., Dupre, B. & Minster, J. F. *Phil. Trans. R. Soc. A* **297**, 447 (1980).
- Davies, G. F. *Nature* **290**, 208 (1981).
- Jambon, A. & Shelby, J. E. *Earth planet. Sci. Lett.* **51**, 206 (1980).
- Kaneoka, I. & Takaoka, N. *Science* **208**, 1366 (1980).
- Kaneoka, I. & Takaoka, N. *Abstr. 5th int. Conf. Geochronology, Cosmochronology, Isotope Geology* (1982).
- Craig, H. & Lupton, J. *Earth planet. Sci. Lett.* **31**, 369 (1976).
- Lupton, J. E., Weiss, R. F. & Craig, H. *Nature* **266**, 244 (1977).
- Kurz, M. D., Jenkins, W. J. & Hart, S. R. *Nature* **297**, 43 (1982).
- Nagao, K., Takaoka, N. & Matsubayashi, O. *Earth planet. Sci. Lett.* **53**, 175 (1981).
- Kyser, T. K. & Rison, W. *J. geophys. Res.* **87**, 5611 (1982).
- Tolstikhin, I. N. in *Terrestrial Rare Gases* (eds Alexander, E. C. Jr & Ozima, M.) (Center for Academic Publishing, Japan, 1978).

Mantle eclogite and carbonate as sources of sodic carbonatites and alkalic magmas

Allan H. Treiman & Eric J. Essene

Department of Geological Sciences, University of Michigan, Ann Arbor, Michigan 48109, USA

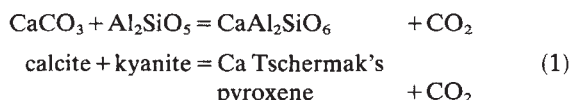
At mantle pressures, partial melts of peridotite + CO₂ are carbonate-rich, but contain insufficient sodium to form most carbonatites and associated silicate magmas. A possible source rock is eclogite, and extrapolations from experiments suggest that eclogite + CO₂ will coexist with melt along geotherms of low heat flow. Thermochemical calculations yield high carbon-

ate activities, and imply that the melt is carbonatitic. The melt can be modelled in the system jadeite–calcite–CO₂ and will, on emplacement, form mineral assemblages characteristic of a wide range of alkalic rocks. The tectonic settings of carbonatites are consistent with an origin as partial melts of eclogite + carbonate + CO₂ under areas of low surface heat flow, and suggest that carbonatite emplacement is not related to mantle hotspots.

At pressures above 25 kbar, partial melts in the system peridotite + CO₂ are rich in carbonate, and have been correlated with kimberlite or carbonatite magmas^{1,2}. However, peridotite does not contain enough sodium to allow formation of most carbonatites and their associated sodic mafic magmas (ijolite, melilitite), which contain sodic minerals or have produced sodic metasomatic aureoles³. The sodium is presumably derived from the mantle, as isotope studies of carbonatites indicate minimal crustal contamination^{4–7}. A possible source of sodium is jadeitic pyroxene (omphacite) in mantle eclogite (a widespread, although minor component of the upper mantle⁸), and we have considered phase relations among eclogite, carbonate, and CO₂ to determine whether carbonatites and their associated silicate rocks might be generated by partial melting of eclogite + carbonate + CO₂.

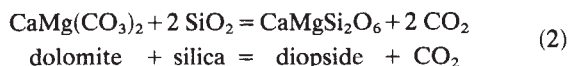
The mantle source of carbonatite melts presumably must contain significant carbonate, but a carbonate mineral (calcite) has been reported only once in mantle eclogite⁹. However, carbonation reactions involving omphacite (jadeite–diopside–Ca Tschermak's pyroxene solid solution), grossular–pyroxene garnet, and minor eclogite minerals (kyanite, corundum¹⁰, or a silica polymorph^{11,12}) can be used to limit the activity of carbonate minerals in equilibrium with eclogite. Phase relations were calculated with the computer program EQUILI (Wall and E.J.E., unpublished), using free energy of CO₂ from Bottinga and Richet¹³, and data for solids phases from the literature^{14–16}.

Among the possible equilibria

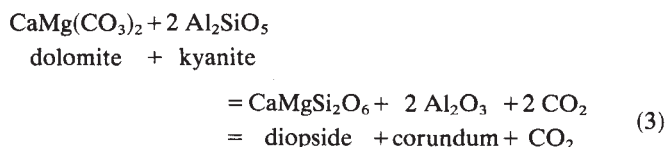


is the simplest and most widely applicable. Omphacite from mantle eclogite¹⁰ generally contains 5% Al(IV) and 60% Ca in the M2 site, so that the activity¹⁷ of Ca Tschermak's pyroxene (CaTs) in a typical mantle eclogite is estimated as $0.05 \times 0.60 = 0.03$. For such a pyroxene, Fig. 1a shows the ratio of calcite activity to carbon dioxide activity buffered by reaction (1). If kyanite is absent (as in most eclogites), the activity ratio isopleths of Fig. 1a represent lower limits, and calcite activity will be higher.

Similarly, carbonation reactions constrain the activity of dolomite in equilibrium with eclogite. The reaction



can be used to constrain the ratio $a(\text{Dol})/a^2(\text{CO}_2)$. Omphacitic pyroxene in mantle eclogite contains significant diopside component¹⁰, equivalent on average to $a(\text{Diop}) = 0.3$ using an ideal ionic activity model. Figure 1b shows $a(\text{Dol})/a^2(\text{CO}_2)$ buffered by such a pyroxene in reaction (2), with the quartz–coesite equilibrium¹⁸. For eclogites lacking quartz or coesite, the activity isopleths of Fig. 1b are minima, and dolomite activity will be higher. For kyanite–corundum eclogites¹⁰, activity ratio $a(\text{Dol})/a^2(\text{CO}_2)$ may be calculated from the reaction



Activity ratio isopleths for reaction (3) are shown in Fig. 1c; for eclogites containing diopside and kyanite, but no corun-

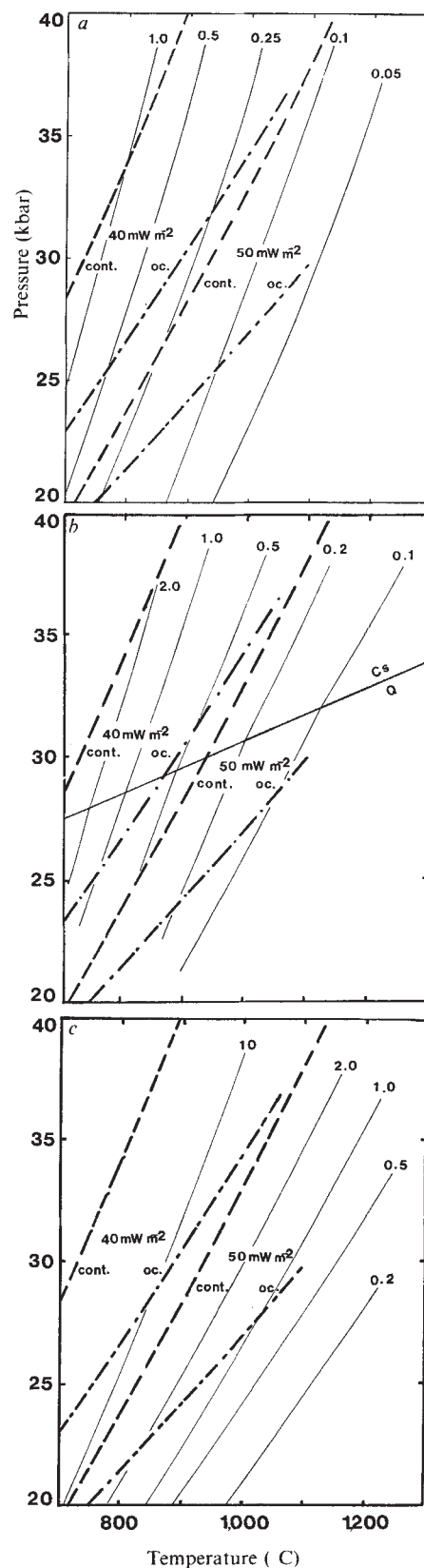


Fig. 1 a, Calculated ratios $a(\text{Cc})/a(\text{CO}_2)$ buffered by reaction (1) with $a(\text{CaTs}) = 0.03$. For kyanite-absent eclogites, plotted values are minima. b, Calculated ratios $a(\text{Dol})/a^2(\text{CO}_2)$ buffered by reaction (2) with $a(\text{Diop}) = 0.3$. For eclogites without quartz or coesite, plotted values are minima. c, Calculated ratios $a(\text{Dol})/a^2(\text{CO}_2)$ buffered by reaction (3) with $a(\text{Diop}) = 0.3$. For eclogites with kyanite but without corundum, plotted values are maxima. Calculated oceanic (oc.) and continental (cont.) geotherms and their surface heat flow are shown as dash-dot lines and dashed lines respectively.

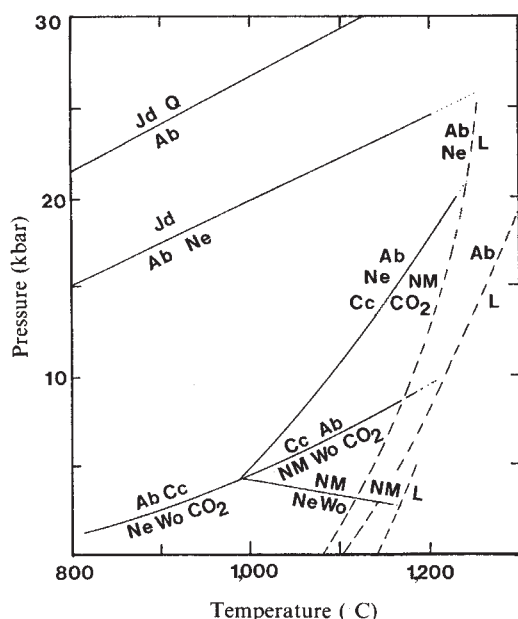


Fig. 2 Calculated phase relations in $\text{NaAlSiO}_4\text{-CaO-SiO}_2\text{-CO}_2$. Melting relations are for the vapour-absent sub-system¹⁷, as solidi for the vapour-saturated system are not known. Metastable reactions are dotted. Formulae and abbreviations for minerals in the system $\text{NaAlSiO}_4\text{-CaO-SiO}_2\text{-CO}_2$ are—calcite: Cc, CaCO_3 ; wollastonite: Wo, CaSiO_3 ; albite: Ab, $\text{NaAlSi}_3\text{O}_8$; jadeite: Jd, $\text{NaAlSi}_2\text{O}_6$; nepheline: Ne, NaAlSiO_4 ; soda-melilite: NM, $\text{CaNaAlSi}_2\text{O}_7$; quartz: Q, SiO_2 ; coesite: Cs, SiO_2 .

dum, the isopleths of Fig. 1c represent upper limits on $a(\text{Dol})/a^2(\text{CO}_2)$.

Essential to application of these equilibria is knowledge of CO_2 fugacities within the mantle. We assume that CO_2 pressure is a significant fraction of load pressure, because fluid inclusions from mantle rocks¹⁹⁻²¹ are generally rich in CO_2 , and experimental studies of carbonatite genesis yield results compatible with significant CO_2 in the source region for carbonatites. For pure CO_2 gas, the isopleths of Fig. 1 show that carbonate activities are significant only below 1,000 °C in the pressure range shown.

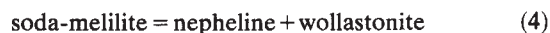
Carbonation reactions in the system $\text{CaO-Al}_2\text{O}_3\text{-SiO}_2\text{-CO}_2$ involving the phases (or components) calcite, grossular, Ca-Tschermak's pyroxene, kyanite, corundum, or quartz-coesite can be used as buffers of CO_2 activity. For the garnet and pyroxene of natural eclogite, calcite activities are similar to those of Fig. 1a. Similarly, in the system containing MgO (including the above minerals plus pyrope, diopside, and dolomite), dolomite activities are buffered to values near those of Fig. 1b and c.

If carbonatites are to form from melting of eclogite + carbonate + CO_2 , eclogite + CO_2 must melt at temperatures which will allow significant carbonate activities in the resultant liquid. Melting relationships of eclogite + CO_2 carbonate are not known, so that extrapolation from similar systems is necessary. Melting relationships in peridotite + CO_2 are well known^{1,2}, and the solidus in that system shows a sharp inflection to lower temperatures at about 25 kbar, with the minimum at about 1,200 °C and 27 kbar. Above that pressure, the solidus temperature increases monotonically. Peridotite + CO_2 is not directly relevant to eclogite, but the solidi in related fluid-absent systems with sodium (the joins jadeite-forsterite and jadeite-enstatite) are ~400 °C below the solidi in similar systems without sodium (diopside-forsterite and diopside-enstatite)²². Addition of garnet to these systems must reduce the solidi, but the effect will be minor because garnet participates little in early partial melting^{23,24}. By extrapolating these results to vapour-saturated systems, one may expect eclogite + CO_2 to

melt near 800 °C at 25 kbar. Similarly, eclogite + H_2O begins to melt near 700 °C at 25 kbar (ref. 25), and H_2O and CO_2 are found to depress melting similarly at this pressure²⁶. Thus, one may expect eclogite + CO_2 to melt between 700 and 1,000 °C at pressures of 25 kbar. Both extrapolations suggest that eclogite + CO_2 should melt within the field of eclogite + carbonate stability, and thus suggest that carbonated eclogite could reasonably be the source of carbonatite magmas.

At pressures greater than 25 kbar, partial melts in the peridotite + CO_2 systems are carbonate-rich^{1,2,27}, and it is reasonable to assume that carbonate-rich melts will also result from partial fusion of eclogite + CO_2 at similar pressures. One must then know if partial melts of eclogite + carbonate + CO_2 can have compositions similar to those of natural carbonatites.

As a model for interactions in the natural system, phase relations in the system $\text{NaAlSiO}_4\text{-CaO-SiO}_2\text{-CO}_2$ have been calculated as above (Fig. 2). That system contains most sodium-bearing minerals relevant to eclogite and carbonatite including soda-melilite ($\text{NaCaAlSi}_2\text{O}_7$), which accounts for up to 50 mol% of igneous melilites²⁸. Solid-solid reactions involving jadeite are from Boettcher and Wyllie²⁹ and Holland³⁰, and the free energy of soda-melilite is from the reaction

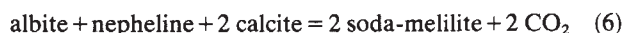


estimated by Yoder²⁸. For clarity, Fig. 2 omits polymorphic transitions and the calcite-quartz-wollastonite- CO_2 equilibrium. Much of Fig. 2 is metastable with respect to melting reactions³¹, but few reactions in the vapour-saturated system are known³².

If melting in the system jadeite-calcite- CO_2 is essentially congruent, the resultant melts will contain significant soda-melilite component by



(noted by Kushiro²⁸), although the reaction is metastable with respect to melting relations. If a partial melt in the system jadeite-calcite- CO_2 is emplaced at the Earth's surface (1 bar pressure), it will crystallize to the mineral assemblage nepheline + wollastonite + calcite found in ijolites and some nephelinite lavas. In natural magmas, melilite may be stabilized to 1 bar by akermanite content, and the assemblages melilite ± calcite ± nepheline ± wollastonite will result. These assemblages are characteristic of melilitite lavas and some carbonatites (like that at Oka, Quebec³³). At pressures greater than those defined by the equilibria



and



crystallization of melts in the system jadeite-calcite- CO_2 will produce calcite + albite ± nepheline ± wollastonite rocks. These assemblages may be compared with the carbonatitic alkalic rocks of the Bancroft-Haliburton area, Canada³, feldspar-bearing carbonatites³, or carbonate albitites³⁴. Water content of the mantle fluid, degree of partial melting, and temperature of formation may control the carbonate content of the melt.

The Na_2O and CO_2 content of the melts cannot be estimated because so little is known of the melting relations of carbonated eclogite, and because the original Na_2O and CO_2 content of carbonatite magmas can be affected by fractional crystallization, loss of volatiles to a fenitizing fluid, or later alteration. Carbonatites of the Oka complex (Quebec, Canada) are among the most sodic known, and include rocks with up to 4.5 wt% Na_2O , which is a least upper bound on possible Na content of carbonatite magmas. Partial melts of uncarbonated eclogite^{23,24} contain between 3 and 9 wt% Na_2O , consistent with the sodium content of the Oka rocks.

To investigate the tectonic implications of a genetic connection between carbonatite and mantle eclogite, calculated carbonate/CO₂ activity ratios (Fig. 1) may be compared with calculated geotherms for both continental and oceanic lithosphere³⁵. The continental lithosphere is modelled as an 8 km thick granitic layer ($\rho = 2.7 \text{ g cm}^{-3}$) and a 32 km thick granulitic layer ($\rho = 3.0 \text{ g cm}^{-3}$) overlying mantle eclogite or peridotite ($\rho = 3.4 \text{ g cm}^{-3}$). Oceanic lithosphere is modelled as a 10-km thick gabbroic layer ($\rho = 2.95 \text{ g cm}^{-3}$) overlying eclogite or peridotite. Continental geotherms of 50 and 40 mW m⁻² approximate geothermal gradients in stable cratons and Precambrian shields respectively. The 50 and 40 mW m⁻² oceanic geotherms approximate lithosphere 90 and greater than 120-Myr old respectively³⁶.

If the mantle is generally saturated with a CO₂-rich fluid¹⁹⁻²¹, kyanite eclogite will buffer calcite activities to moderately high values along both plotted continental geotherms and the lower heat flow oceanic geotherm. If kyanite and quartz or coesite are absent, the isopleths of Fig. 1 represent lower limits on carbonate phase activities. Thus, mantle eclogite under areas of low heat flow should be in equilibrium with calcic dolomite or a carbonate-rich fluid or melt, and be a potential source for carbonatites. In fact, carbonatites on continents are restricted to shields and cratons³, and a few are reported from the Cape Verde Islands, presumably on old oceanic crust³⁷, so that calculated-phase equilibria are compatible with the tectonic setting of carbonatites.

Because carbonatite melts are stable with mantle eclogite along normal geotherms corresponding to low surface heat flow, no unusual source of heat is needed to generate carbonatites. In fact, carbonatite melts cannot be stable with eclogite along the high heat-flow geotherms normally ascribed to 'hotspots' or mantle thermal anomalies (Fig. 1). This suggests that carbonatites should not be used as hotspot tracers (as has been done with kimberlites³⁸), but that carbonatite vulcanism can be reasonably ascribed to activation of deep crustal fractures³⁹.

We thank Dr Henry N. Pollack for discussions of heat flow. This research was supported in part by NSF grant EAR-7923664. Contribution 387 from the Mineralogical Laboratory, Department of Geological Sciences, University of Michigan.

Received 23 June 1982; accepted 16 February 1983.

- Eggler, D. H. & Holloway, J. R. *Oreg. Dep. Geol. Miner. Ind. Bull.* **96**, 15-36 (1977).
- Wyllie, P. J. *Am. Miner.* **64**, 469-500 (1979).
- Heinrich, E. W. *The Geology of Carbonatites* (Rand-McNally, Chicago, 1966).
- Deines, P. & Gold, D. P. *Geochim. cosmochim. Acta* **37**, 1709-1733 (1973).
- Suwa, K., Oana, S., Wada, H. & Osaki, S. *Phys. Chem. Earth* **9**, 735-745 (1975).
- Rock, N. M. S. *Contr. Miner. Petrol.* **56**, 205-228 (1976).
- Basu, A. R. & Tatsumoto, M. *Contr. Miner. Petrol.* **75**, 43-54 (1980).
- Ringwood, A. E. *Composition and Petrology of the Earth's Mantle* (Wiley, New York, 1975).
- Hunter, W. C. & Smith, D. *Contr. Miner. Petrol.* **76**, 312-320 (1980).
- Sobolev, N. V. *Deep Seated Inclusions in Kimberlites and the Problem of the Composition of the Upper Mantle* (American Geophysical Union, Washington DC, 1977).
- Smyth, J. R. & Hatton, C. J. *Earth planet. Sci. Lett.* **34**, 284-290 (1977).
- Ponomarenko, A. I., Spetsius, Z. V. & Lyubushkin, V. A. *Dokl. Akad. Nauk SSSR* **236**, 126-128 (1977).
- Bottinga, Y. & Richet, P. *Am. J. Sci.* **281**, 615-660 (1981).
- Clark, S. P. *Geol. Soc. Am. Mem.* **97** (1966).
- Robie, R. A., Hemingway, B. S. & Fisher, J. R. *Bull. U.S. geol. Surv.* **1452**, (1979).
- Haas, J. L. Jr., Robinson, G. L. Jr. & Hemingway, B. S. *J. phys. Chem. Ref. Data* **10**, 575-669 (1981).
- Wood, B. J. *Am. Miner.* **61**, 599-602 (1976).
- Mirwald, P. W. & Massonne, H.-J. *J. geophys. Res.* **85B**, 6983-6990 (1980).
- Roedder, E. *Am. Miner.* **50**, 1746-1782 (1965).
- Green, H. W. & Radcliffe, S. V. *Bull. geol. Soc. Am.* **86**, 846-852 (1975).
- Murck, B. W., Burruss, R. C. & Hollister, L. C. *Am. Miner.* **63**, 40-46 (1978).
- Windom, K. E. & Boettcher, A. L. *Am. J. Sci.* **281**, 335-351 (1981).
- Switzer, G. & Melson, W. G. *Smithsonian Contr. Earth Sci.* **1**, 1-9 (1969).
- Windom, K. E. & Boettcher, A. L. *J. Geol.* **88**, 705-712 (1975).
- Lambert, I. B. & Wyllie, P. J. *Science* **169**, 764-766 (1970).
- Wyllie, P. J. *J. Geol.* **86**, 687-713 (1978).
- Wyllie, P. J. & Huang, W. L. *Geology* **3**, 621-624 (1975).
- Yoder, H. S. Jr *Fortschr. Miner.* **50**, 140-173 (1973).
- Boettcher, A. L. & Wyllie, P. J. *Geochim. cosmochim. Acta* **32**, 999-1012 (1968).
- Holland, T. *Am. Miner.* **65**, 129-134 (1980).
- Bell, P. M. & Roseboom, E. H. Jr *Miner. Soc. Am. Spec. Pap.* **2**, 151-161 (1969).
- Eggler, D. H. & Kadik, A. A. *Am. Miner.* **64**, 1036-1048 (1979).
- Gold, D. P., Vallee, M. & Charette, J.-P. *Bull. Can. Min. Metall.* **60**, 1131-1144 (1967).
- Shimron, A. E. *Min. Mag.* **40**, 13-24 (1975).
- Chapman, D. S. & Pollack, H. N. *Geology* **5**, 265-268 (1977).
- Parsons, B. & Selater, J. G. *J. geophys. Res.* **82**, 803-827 (1977).
- Silva, L. C., Le Bas, M. J. & Robertson, A. H. F. *Nature* **294**, 644-645 (1981).
- Crough, S. T., Morgan, W. J. & Hargraves, R. B. *Earth planet. Sci. Lett.* **50**, 260-274 (1980).
- Bailey, D. K. *J. geol. Soc. Lond.* **133**, 103-106 (1977).

Rapid development of tolerance to the behavioural actions of cholecystokinin

J. N. Crawley*

Neurobiology Program, Central Research and Development Department, E. I. du Pont de Nemours and Company, Glenolden, Pennsylvania 19036, USA

M. C. Beinfeld

Department of Pharmacology, University of St Louis School of Medicine, St Louis, Missouri 63104, USA

Cholecystokinin (CCK) acts acutely to inhibit food consumption in fasted rats¹⁻⁴, mice⁵⁻⁷, sheep⁸, pigs⁹, monkeys¹⁰ and humans¹¹⁻¹³. CCK has been proposed as a satiety signal, inducing the behavioural sequence of satiety², or as an aversive internal stimulus, which inhibits food intake by inducing malaise⁴. Reductions in food intake and related exploratory behaviours are initiated by CCK at its peripheral receptor in the gut, which appears to transmit sensory feedback via the vagus nerve to brain regions mediating appetitive behaviours¹⁵⁻¹⁷. The therapeutic potential of CCK as an appetite suppressant in obesity syndromes rests on the demonstration of significant, long-lasting body weight reduction. Chronic CCK administration by repeated injections is problematic, since this peptide is rapidly degraded *in vivo*. We chose the Alzet constant infusion osmotic minipump to investigate possible alterations in body weight and food intake during continuous infusion of CCK. We now report that no change was detected in either body weight or total daily food consumption at any time point during 2 weeks of intraperitoneally (i.p.) infused CCK. The mechanism underlying the lack of chronic CCK effects appears to be a rapid development of behavioural tolerance. Acute challenge doses of CCK which induced satiety-related behaviours in saline-infused rats were ineffective in CCK-infused rats. The behavioural tolerance was apparent within a few hours of minipump implantation. These results provide the first evidence that rapid and reversible tolerance develops to the actions of a gut peptide.

Male Sprague-Dawley rats (200-250 g starting weight), individually housed in a temperature and light-controlled environment, were maintained on an *ad libitum* diet of Purina rat chow and water or condensed milk enriched with multivitamins (Poly-Vi-Sol, Mead Johnson). Alzet osmotic minipumps (Model 2001 $1 \mu\text{l h}^{-1}$, 7 days or Model 2002, $0.5 \mu\text{l h}^{-1}$, 14 days) were surgically implanted in the abdominal cavity under halothane anaesthesia. Pumps were implanted between hours 4 and 7 of the 12-h lights-on period, when food consumption is normally low, but the gut contains food recently consumed. Fourteen-day pumps were filled with saline or CCK8-sulphate (Bachem) to deliver 0, 0.1 or $1.0 \mu\text{g per kg body weight per h}$ in experiment 1. Seven-day pumps were filled with saline or CCK8-sulphate to deliver 0, 0.05, or $1.0 \mu\text{g per kg per h}$ in experiment 2. Filled pumps were pre-warmed to ensure immediate, uniform delivery at the time of implantation. Stability of CCK in the implanted pumps was analysed by HPLC and radioimmunoassay^{18,19} of fluid removed at day 13 from the 14-day pumps. Each animal was weighed daily. Food consumption was recorded in experiment 1 by daily weighing of the rat chow container and in experiment 2 by volumetric measuring of the liquid diet from calibrated bottles which minimized spillage. In experiment 2, satiety-related reductions in exploratory behaviours^{5,20,21} were analysed in 5-min sessions in a video-tracking computer-assisted animal behaviour monitor²², immediately after challenge doses of CCK $5 \mu\text{g per kg i.p.}$ at 4 h, 1 and 4 days after pump implantation and 1 day after the pump contents were depleted. This monitoring system records the location of the animal as an x-y

* Present address: Unit of Behavioral Neuropharmacology, Clinical Neuroscience Branch, National Institute of Mental Health, Building 10, Room 4N214, Bethesda, Maryland 20205, USA.

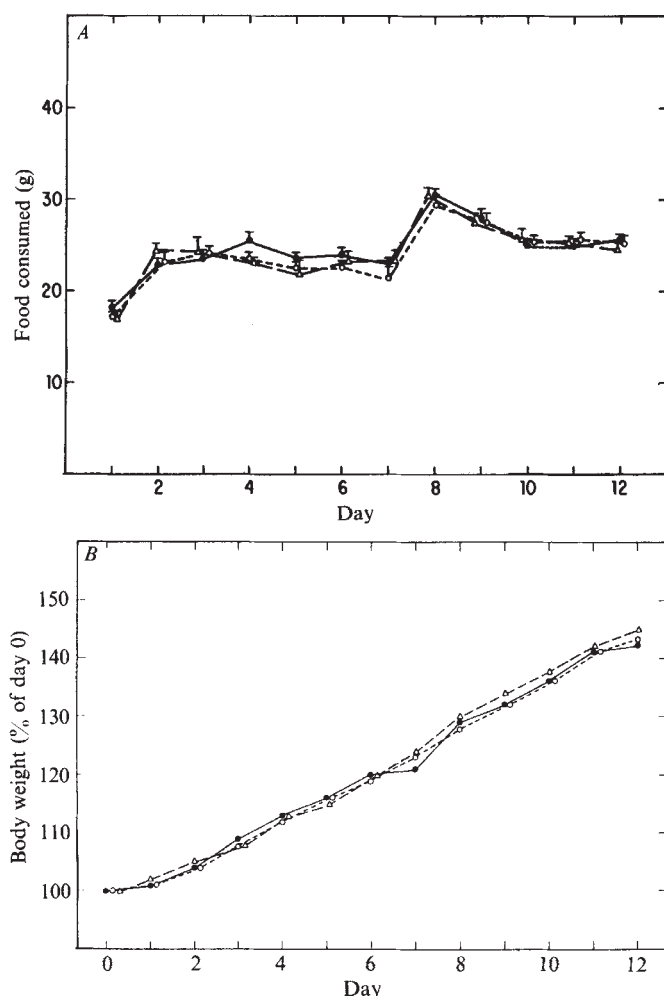


Fig. 1 Food consumption (A) and body weight (B) were measured daily in rats implanted intraperitoneally with Alzet osmotic minipumps containing CCK8-sulphate 1.0 µg per kg per µl (Δ), 0.1 µg per kg per µl ($\circ$), or saline ($\bullet$). Chronic CCK had no effect on daily food intake ($F_{41,378} = 0.39$, NS) or body weight ($F_{41,378} = 0.75$, NS) as compared with saline controls. Values are expressed as mean $\pm$ s.e.m. for A, and as mean per cent of starting weight for B, where variability around the mean was less than 5%.

coordinate every 0.1 s. Parameters of movement, pauses, and exploration of an unfamiliar 60 $\times$ 60 cm open field, containing a novel wire mesh box, were analysed from the raw data by the OMNI program on a VAX-4 computer. One way analysis of variance (ANOVA) was used to test significance of treatment effects, with comparisons of individual means by the Neuman-Keuls *a posteriori* test.

Body weight and food consumption in CCK-infused rats remained identical to saline controls throughout the 14-day infusion period (Fig. 1). Seven days of infusion yielded results similar to 14 days of infusion. Body weights and total daily food consumption for rats infused with 1.0 µg/kg per h or 0.05 µg/kg per h CCK were not significantly different from saline-infused control rats at any time point during or after the 7-day infusion ($F_{20,198} = 0.71$, not significant (NS), for body weight; $F_{20,198} = 1.23$, NS, for food intake).

HPLC reverse phase chromatography of the pump contents showed a peak of optical density eluting at a retention time identical to the CCK8-sulphate standard (Fig. 2). Radioimmunoassay of this peak recovered 0.0058 µg µl⁻¹ CCK8-sulphate from the 0.1 µg/kg per µl-filled pumps and 0.029 µg µl⁻¹ CCK8-sulphate from the 1.0 µg/kg per µl-filled pumps, and respectively 23% and 12% of the concentrations originally implanted.

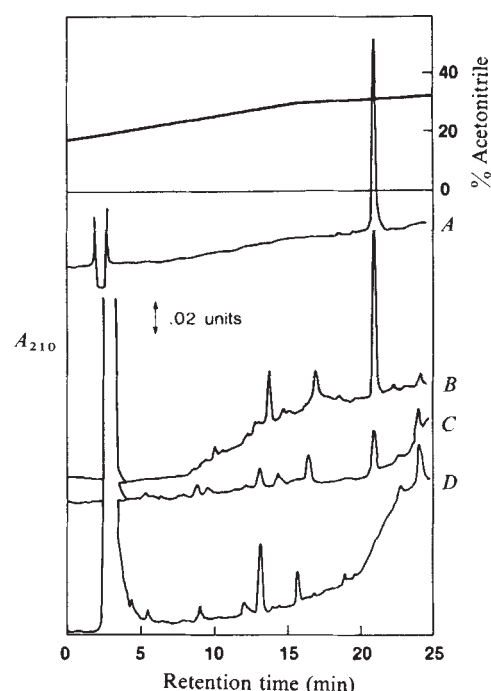


Fig. 2 HPLC reverse phase chromatography of CCK8-sulphate standard (A) and the contents of Alzet minipumps filled with 1.0 µg per kg per µl CCK8-sulphate (C), or saline (D). Pump contents were obtained by aspirating the liquid from hemisectioned pumps removed from the peritoneal cavity of rats 13 days after implantation of 14-day pumps. Samples were run on a Supelco C18 column, 0.46 $\times$ 25 cm, in triethylamine phosphoric acid buffer, pH 3.25, with a nonlinear gradient of 20–32% acetonitrile. Radioimmunoassay for CCK verified authentic CCK8-sulphate in the 21-min peak. Extra peaks may represent endogenous substances diffusing into the pump.

Acute challenge doses of CCK decreased satiety-related exploratory parameters as previously described^{5,20,21}, in rats whose pumps contained saline (Fig. 3). No response to the challenge dose of CCK was detected in rats whose pumps contained either 1.0 µg/kg per h or 0.05 µg/kg per h of CCK. The blockade of responsiveness was significant at both doses at 4 h, 1 day, and 4 days after pump implantation. A trend towards greater tolerance at the higher dose was not significant. One day after the pump contents were expended, all rats showed the normal reduction in exploratory behaviours to a challenge dose of CCK.

The attenuation of behavioural response to a challenge dose of CCK within 4 h after the start of constant CCK infusion suggests that the earliest time point at which tolerance appears may be less than 4 h. Minipump implantation is not a useful technique for the time period of 0–4 h, since recovery from the acute effects of surgery and halothane anaesthesia requires at least 1 h. To address this question, a separate group of naive, unimplanted rats were treated with a single acute i.p. dose of 5.0 µg/kg CCK, followed by a second, challenge, dose of 5.0 µg/kg CCK at 0, 30, 60, or 120 min after the first. Immediately following the second dose, each animal was tested for 5 min as above. Since the half life of CCK in plasma is approximately 3 min, very little CCK from the first injection should remain at the time of the second injection. Table 1 illustrates the significant attenuation of CCK effects when the second injection was 1 or 2 h after the first.

Insensitivity to the challenge dose of CCK within hours after the start of constant CCK infusion or after an acute injection suggests that tolerance had developed to the behavioural effects of CCK. The concentration of infused CCK required for the development of behavioural tolerance appears to be very low, since even the lower concentration of 0.05 µg/kg per h, which was partially degraded internally, appeared to block responses to the CCK challenge dose. The tolerance is quickly

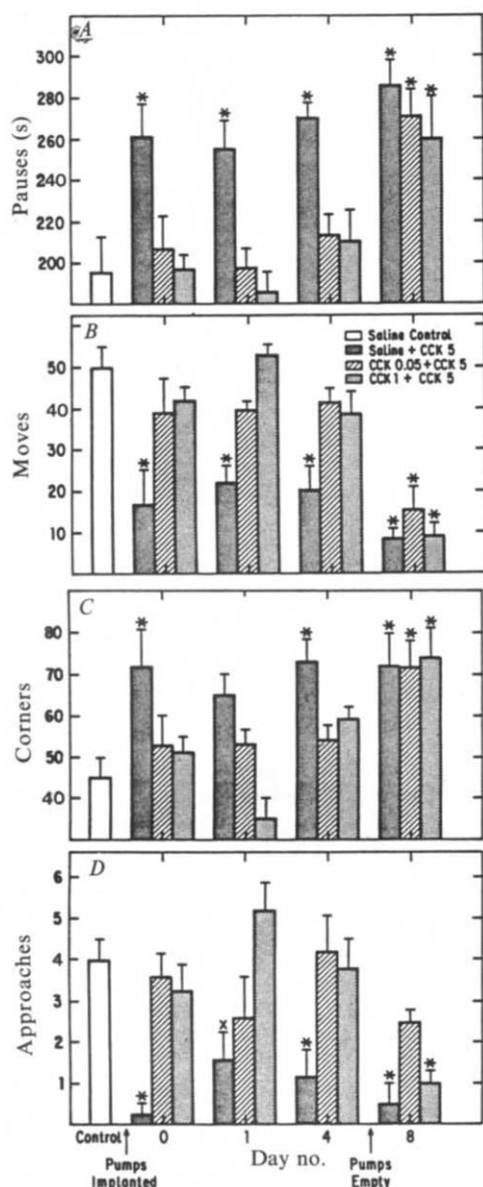


Fig. 3 Reductions in exploratory behaviours representative of the satiety syndrome were produced by an acute challenge dose of CCK8-sulphate 5 μg per kg i.p. in rats whose osmotic infusion pump contained saline. Rats constantly infused with either the low or the high concentration of CCK did not respond to the challenge dose of CCK. **A**, Total pause time, defined as cumulative time spent in episodes of pausing, in which the rat remained immobile within a 5-cm radius for a minimum of 3 s: $F_{11,108} = 5.08$, $P < 0.01$. **B**, Number of movements, defined as total number of locomotor moves initiated: $F_{11,108} = 9.26$, $P < 0.01$. Per cent time spent in corners, defined as percentage of total session time in which the rat was within 3 cm of a corner of the arena: $F_{11,108} = 2.10$, $P < 0.05$. **D**, Number of approaches to novel object, defined as the total number of times the rat was in physical contact with the wire mesh box for a minimum of 3 s: $F_{11,108} = 2.64$, $P < 0.01$. The lack of response to the CCK challenge dose was detectable within 4 h after pump implantation. Response to the acute CCK challenge returned to normal within 1 day after the contents of the pump were expended. Values are expressed as mean $\pm$ s.e.m. $n = 10$ for each group. * $P < 0.05$ by Neuman-Keuls comparison of individual means.

reversed, without rebound, as normal response to the CCK challenge dose were restored within 1 day after the 7-day pumps were depleted.

The rapid development of tolerance to the behavioural actions of CCK provides one explanation for the inability of chronic CCK to reduce body weight. This possibility was suggested by inconsistent feeding responses to repeated daily CCK

Table 1 Time course of tolerance to repeated CCK injections

Time between injections (min)	Total pause time (s)	No. of movements	% Time spent in corners	No. of approaches to novel objects
0	245 $\pm$ 13	25.2 $\pm$ 3	73.7 $\pm$ 5	1.0 $\pm$ 1
30	221.1 $\pm$ 9	24.3 $\pm$ 6	65.6 $\pm$ 9	2.8 $\pm$ 1
60	204.0 $\pm$ 13	34.4 $\pm$ 4*	53.4 $\pm$ 9*	3.4 $\pm$ 3
120	179.6 $\pm$ 12*	42.9 $\pm$ 3*	44.9 $\pm$ 5*	7.5 $\pm$ 2*

An acute i.p. injection of 5 μg per kg CCK8 was followed by a second i.p. injection of 5 μg per kg CCK8 at either 0, 30, 60, or 120 min after the first dose. Immediately after the second dose, the animals were tested in a 5-min session using the video-tracking computer-assisted animal behaviour monitor as described in Fig. 3 legend. Significant attenuation of the behavioural effects of CCK began between 1 and 2 h after acute injection. Values are expressed as mean $\pm$ s.e.m. $n = 8$ for each treatment group. Pauses: $F_{3,28} = 2.96$, $P < 0.05$; no. of movements: $F_{3,28} = 3.65$, $P < 0.05$; per cent time in corners: $F_{3,28} = 3.09$, $P < 0.05$; no. of approaches: $F_{3,28} = 3.23$, $P < 0.05$. * $P < 0.05$ by Newman-Keuls comparison of individual means.

injections, reported earlier by Mineka and Snowden²⁴. CCK or its analogues will prove clinically useful only if the development of tolerance can be by-passed.

Several mechanisms might explain the observed behavioural tolerance to CCK. If CCK actions are mediated by gastric afferents¹⁵⁻¹⁷, CCK may inhibit gastric emptying and intensify a food-dependent distension²⁴. Tolerance could develop by rapid extinction of the conditioned satiety component of the distension²⁵. The mechanisms of action could be the interaction between CCK and food remaining in the digestive tract since the last meal.

Induction of metabolic enzymes which degrade peptides could increase the rate of degradation of exogenously administered CCK to prevent its action. Trypsin inhibitors which increase serum CCK levels have been shown to decrease meal size²⁷, suggesting that, conversely, increased metabolism might prevent CCK from producing satiety effects.

Another mechanism which could produce the observed tolerance is the development of CCK receptor subsensitivity. Rapid change in receptor number is a well established phenomenon in many neurotransmitter systems²⁷⁻³⁶. Gut peptides may similarly induce rapid, sensitive, regulatory changes in receptor function, at the level of the gut, or possibly on the CCK receptors recently found on the vagus nerve³⁷.

Other sites of tolerance development could be intracellular second messengers. Gardner *et al.*³⁸ report desensitization of physiological responsiveness in pancreatic acinar cells to repeated administration of gut peptides. Dispersed acini preincubated with bombesin or CCK secreted 60-75% less pancreatic amylase when resuspended and incubated for a second time with these peptides. Desensitization was detected within 5 min and was maximal when 40-60 min were allowed between the two CCK injections. Desensitization to the physiological effects of CCK may occur at some step after receptor occupation, since the response to other agents which mobilize cellular calcium is also attenuated by CCK preincubation. Desensitization of a second messenger such as phosphatidylinositol may be critical to both the physiological and behavioural actions of CCK.

The present study does not directly address the controversial issue of whether CCK inhibits feeding behaviours by inducing feelings of satiety or feelings of sickness. Reductions in exploratory parameters used above simply provide a useful index of CCK actions for the present investigation of tolerance. The primary cause of the reduced exploration could be either satiety or malaise.

This research was supported in part by NIH grant NS 18335 and a grant from the American Parkinson Disease Foundation. We thank Lisa Blumstein and Beth Double for technical assistance.

Received 29 September 1982; accepted 11 February 1983.

- Gibbs, R. A., Young, C. & Smith, G. P. *J. comp. Physiol. Psychol.* **84**, 488-495 (1973).
- Antin, J., Gibbs, J., Holt, J., Young, R. C. & Smith, G. P. *J. comp. Physiol. Psychol.* **89**, 784-790 (1975).
- Hsiao, S., Wang, C. H. & Schallert, T. *Physiol. Behav.* **23**, 909-914 (1979).
- McLaughlin, C. & Baile, C. A. *Physiol. Behav.* **25**, 543-548 (1980).
- Crawley, J. N., Hays, S. E., Paul, S. M. & Goodwin, F. K. *Physiol. Behav.* **27**, 407-411 (1981).
- Koopmans, H. S., Deutch, J. A. & Branson, P. J. *Behav. Biol.* **7**, 441-444 (1972).
- Parrott, R. F. & Batt, R. A. L. *Physiol. Behav.* **24**, 751-753 (1980).
- Della-Fera, M. A. & Baile, C. A. *Science* **206**, 471-473 (1979).
- Parrott, R. F. & Baldwin, B. A. *Physiol. Behav.* **24**, 419-422 (1980).
- Falasco, J. D., Smith, G. P. & Gibbs, J. *Physiol. Behav.* **23**, 887-890 (1979).
- Stacher, G., Bauer, H. & Steininger, H. *Physiol. Behav.* **23**, 325-331 (1979).
- Sturdevant, R. A. L. & Goetz, H. *Nature* **261**, 713-715 (1976).
- Kissileff, H. R., Pi-Sunyer, F. X., Thornton, J. & Smith, G. P. *Am. J. clin. Nutr.* **34**, 154-160 (1981).
- Deutsch, J. A. & Hardy, W. T. *Nature* **266**, 196 (1977).
- Smith, G. P., Jerome, C., Cushin, B. J., Eterno, R. & Simansky, K. J. *Science* **213**, 1036-1037 (1981).
- Crawley, J. N., Hays, S. E. & Paul, S. M. *Eur. J. Pharmacol.* **73**, 379-380 (1981).
- Morley, J. E., Levine, A. S., Kneip, J. & Grace, M. *Life Sci.* **30**, 1943-1947 (1982).
- Beinfeld, M. C. *Neuropeptides* **1**, 203-209 (1981).
- Beinfeld, M. C., Meyer, D. K., Eskay, R. L., Jensen, R. T. & Brownstein, M. J. *Brain Res.* **212**, 51-57 (1981).
- Crawley, J. N., Hays, S. E., O'Donohue, T. L. & Goodwin, F. K. *Peptides* **2**, Suppl. 1, 123-129 (1981).
- Crawley, J. N., Rojas-Ramirez, J. A. & Mendelson, W. B. *Peptides* **3**, 535-538 (1982).
- Crawley, J. N., Szara, S., Pryor, G. T., Creveling, C. R. & Bernard, B. K. *J. Neurosci. Meth.* **5**, 235-247 (1982).
- Mineka, S. & Snowdon, C. T. *Physiol. Behav.* **21**, 65-72 (1978).
- Moran, T. & McHugh, P. R. *Am. J. Physiol.* **242**, R491-R497 (1982).
- Booth, D. A. *J. comp. Physiol. Psychol.* **81**, 457-471 (1972).
- McLaughlin, C. L., Peikin, S. R. & Baile, C. A. *Am. J. Physiol.* (in press).
- Wolfe, B. B., Harden, T. K., Sporn, J. R. & Molinoff, P. B. *J. Pharmac. exp. Ther.* **207**, 446-457 (1978).
- Mukherjee, C., Caro, J. G. & Lefkowitz, R. J. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1945-1949 (1975).
- Ehlert, F. J., Kokka, N. & Fairhurst, A. S. *Molec. Pharmacol.* **17**, 24-30 (1980).
- Rosenberg, H. C. & Chui, T. H. *Life Sci.* **24**, 803-808 (1979).
- Crawley, J. N., Marangos, P. J., Stivers, J. & Goodwin, F. K. *Neuropharmacology* **21**, 85-90 (1982).
- Lippa, A. S. *et al. Pharmac. Biochem. Behav.* **9**, 853-856 (1978).
- Fagni, L., Gaudry, M. & Lynch, G. *Soc. Neurosci.* **8** (1982).
- Lesniak, M. A. & Roth, J. J. *biol. Chem.* **251**, 3720-3729 (1981).
- Davis, M. E., Akera, T. & Brody, T. M. *J. Pharmac. exp. Ther.* **211**, 112-119 (1979).
- Elazar, Z., Motles, E., Fly, Y. & Simantov, R. *Life Sci.* **24**, 541-548 (1979).
- Zarbin, M. A., Walmsley, J. K., Innis, R. B. & Kuhar, M. J. *Life Sci.* **29**, 697-705 (1981).
- Gardner, J. D., Abdelmoumene, S., Lee, P. C., Collins, S. M. & Jensen, R. T. in *Biology of Normal and Cancerous Exocrine Pancreatic Cells* (eds Ribet, A., Pradayrol, L. & Susini, C.) 165-169 (Elsevier, Amsterdam, 1980).

Stimulation and inhibition of adenylyl cyclases mediated by distinct regulatory proteins

John D. Hildebrandt*, Ronald D. Sekura†, Juan Codina*, Ravi Iyengar*, Charles R. Manclark† & Lutz Birnbaumer*

* Department of Cell Biology, Baylor College of Medicine, Houston, Texas 77030, USA

† Pertussis Branch, Division of Bacterial Products, National Center from Drugs and Biologics, Food and Drug Administration, Bethesda, Maryland 20205, USA

Adenylyl cyclases are under positive and negative control by guanine nucleotides and hormones¹⁻³. Stimulatory responses are mediated by a guanine nucleotide- and Mg-binding regulatory component⁴⁻⁶ (N_s), a protein that has been purified to homogeneity⁷⁻⁹. Inhibitory responses have been hypothesized^{1-3,10,11} to be mediated by an analogous regulatory component (N_i) distinct from N_s , but definitive proof for this is lacking and these effects may result from modulation of N_s activity. Recently, *Bordetella pertussis* toxin has been shown¹² to ADP-ribosylate a peptide that is not part of N_s , and this coincides with attenuation of hormonal inhibition of adenylyl cyclase¹²⁻¹⁷. We show here that *cyc*⁻ S49 cells contain a substrate for ADP-ribosylation by pertussis toxin and that the toxin alters GTP dependent inhibition of *cyc*⁻ adenylyl cyclase activity¹⁸. As *cyc*⁻ S49 cells do not contain N_s by several criteria^{5,19-24}, we conclude that N_i is a distinct and separate regulatory component of adenylyl cyclase.

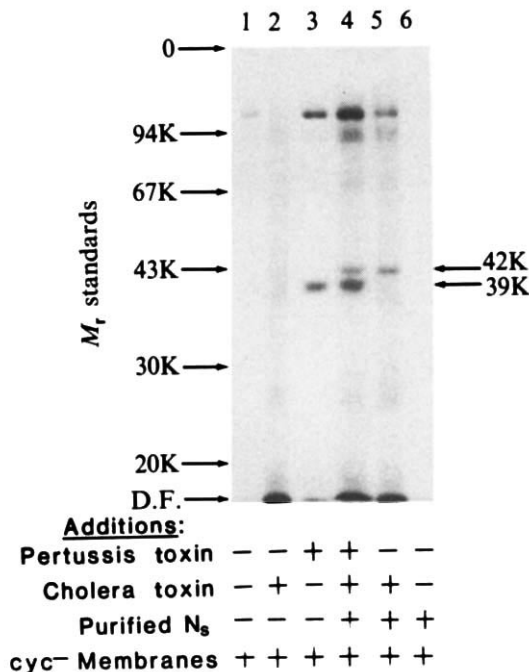


Fig. 1 Incorporation of [³²P]ADP-ribose into *cyc*⁻ S49 cell membrane proteins and purified human erythrocyte N_s . The additions to each sample are indicated. Lane 4 contains a mixture of the samples in lanes 3 and 5 and is included to show that the pertussis toxin substrate in *cyc*⁻ cell membranes is different from the cholera toxin substrate that is part of the purified N_s . *cyc*⁻ S49 cells from the M3B1 strain were grown according to the methods of Bourne *et al.*³⁶ and purified membranes prepared by the procedure of Ross *et al.*²⁴, except that Mg-free buffers were used throughout. Membranes (200 µg) were incubated in a final volume of 100 µl containing 1 mM ATP, 0.5 mM GTP, 15 mM thymidine, 100 mM potassium phosphate pH 7.5, 5 mM dithiothreitol (DTT) and 20 µM [³²P]NAD⁺ (specific activity 500 mCi mmol⁻¹). Incubations were for 30 min at 32 °C. Samples were then diluted fivefold with 10 mM Tris-HCl pH 7.5, 1 mM EDTA and the membranes precipitated by centrifugation at 23,000 r.p.m. for 60 min. The supernatant was removed and the membranes were resuspended in Laemmli's sample buffer³⁷ with 5% β-mercaptoethanol and allowed to stand at room temperature overnight. Samples were electrophoresed in 10% SDS-polyacrylamide gels made by the procedures of Laemmli³⁷. An autoradiogram of the dried gel is shown. Concentrations, when present, were cholera toxin (Sigma), 20 µg ml⁻¹; pertussis toxin, 10 µg ml⁻¹; and purified erythrocyte N_s , 0.5 µg ml⁻¹. The N_s preparation was about 95% pure and consisted of two peptides of M_r 42,000 and 35,000. It was stored in 0.1% lubrol, 20 mM β-mercaptoethanol, 50 mM NaCl, 10 mM Na-HEPES pH 7.5, 30% ethylene glycol at -70 °C and diluted 30-fold for the assay. *B. pertussis* toxin was prepared from cultures of strain 114 (ref. 39) by a modification (R.D.S. and C.R.M., in preparation) of published procedures^{39,40}. Pertussis toxin was preactivated by incubating it in 20 mM DTT for 30 min at 32 °C. Preactivation of cholera toxin is described elsewhere¹⁸.

Recently, we have shown that the *cyc*⁻ adenylyl cyclase is inhibited by GTP, other guanine nucleotides and NaF¹⁸. There was no information in this previous study, however, to suggest that these effects are mediated by the hypothetical N_i thought to confer hormonal inhibition to adenylyl cyclase in other systems¹⁻³. These effects could, in fact, be due to a structurally altered N_s which functionally acts as an inhibitor of catalytic activity. To determine whether or not *cyc*⁻ contains an N_i , we have taken advantage of the recent observation that a toxin isolated from *B. pertussis* blocks hormone-mediated inhibition of adenylyl cyclase activity¹²⁻¹⁷ and at the same time ADP-ribosylates a peptide of molecular weight (M_r) ~40,000 (40K) distinct from the N_s subunit which is a substrate for cholera toxin¹² (There are actually several toxins produced by *B. pertussis*. The one used in these studies, and referred to as pertussis toxin, is that characterized by Ui and collaborators¹²⁻¹⁴ and referred to by them as islet activating protein.) At present,

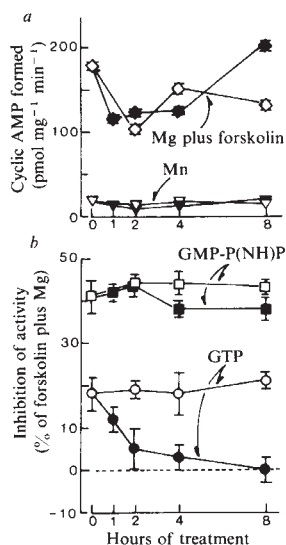


Fig. 2 Effect of treatment of *cyc*⁻ S49 cells with pertussis toxin on their adenylyl cyclase activity. *cyc*⁻ S49 cells were resuspended in fresh culture medium³⁶ at a density of 9.2×10^6 cells ml⁻¹ without (control; open symbols) or with $0.1 \mu\text{g ml}^{-1}$ pertussis toxin (solid symbols). By 8 h of incubation control cultures had increased to 10.7×10^6 cells ml⁻¹ and toxin-treated cultures to 9.7×10^6 cells ml⁻¹. At the indicated times membranes were prepared as described in Fig. 1 legend, except that the procedure was stopped after obtaining the 43,000g pellet. Adenylyl cyclase assays were done in a final volume of 50 μl containing $\sim 10 \mu\text{g}$ membrane protein, 0.1 mM ATP (5×10^6 c.p.m. [α -³²P]ATP), 10 mM MnCl₂ or MgCl₂, 25 mM Na-HEPES pH 8.0, 1 mM EDTA, 1 mM cyclic AMP (10,000 c.p.m. [³H]cAMP) and a nucleoside triphosphate regenerating system composed of 20 mM creatine phosphate, 0.2 mg ml⁻¹ creatine phosphokinase and 0.02 mg ml⁻¹ myokinase. When present, forskolin (Calbiochem) was 100 μM , GTP 10 μM , and GMP-P(NH)P 10 μM . All other procedures were as described recently¹⁸. In *a* values are mean $\pm$ s.d. of three determinations. In *b*, per cent inhibition was calculated as $100(1 - F/T)$ where *F* is the mean of triplicate determinations in the presence of forskolin alone and *T* is that for forskolin plus GTP or GMP-P(NH)P. The error bars in *b* represent s.d.s derived from the following equation⁴¹:

$$\text{s.d.} = \left(\frac{T}{C} \right) \sqrt{\left(\frac{\text{s.d.}_C}{C} \right)^2 + \left(\frac{\text{s.d.}_T}{T} \right)^2}$$

the function of the 40K peptide is unclear, but it does not seem to be a component of purified N_s (refs 7–9). It is either a component of N_i, if this does in fact exist, and interacts directly with the catalytic component of the system, or it is a regulatory element that modulates the activity of N_s and/or inhibitory receptors without interacting directly with the catalytic component of the enzyme. In the experiments presented here, we reasoned that if N_i is distinct from N_s, and if its presence in *cyc*⁻ membranes is the reason for guanine nucleotide regulation

of *cyc*⁻ adenylyl cyclase activity¹⁸, these membranes should contain a pertussis toxin substrate. More importantly, however, treatment of *cyc*⁻ cells with pertussis toxin should alter the pattern of inhibitory responses previously described¹⁸. On the other hand, if *cyc*⁻ membranes contain an altered N_s, these cells may or may not contain a pertussis toxin substrate, but treatment of *cyc*⁻ cells with toxin would not affect the response of *cyc*⁻ adenylyl cyclase to guanine nucleotides.

Figure 1 shows an autoradiogram from an experiment in which *cyc*⁻ S49 cell membranes were incubated with ³²P-NAD⁺ and various combinations of cholera toxin, pertussis toxin and purified N_s from human erythrocytes (J.C., J.D.H., R.I., J. Bryan and L.B., in preparation). With ³²P-NAD⁺ alone only a few high molecular weight proteins were labelled (Fig. 1, lane 1). In the presence of cholera toxin, there was nonspecific labelling of all the major protein bands (Fig. 1, lane 2), but no specific labelling of a 42–45K peptide characteristic of almost all intact adenylyl cyclase systems, including adipocytes²⁵, wild-type S49 cells²², human²⁶, frog²⁷, pigeon^{28,29}, rat³⁰ and turkey³¹ erythrocytes, and liver cells³². This finding is in agreement with the lack of an effect of cholera toxin on inhibition of *cyc*⁻ S49 adenylyl cyclase by guanine nucleotides¹⁸. Although incubation of *cyc*⁻ cell membranes with purified N_s alone (Fig. 1, lane 6) did not substantially alter the incorporation of ³²P-NAD⁺, the presence of purified N_s and cholera toxin together (lane 5) resulted in the specific labelling of a 42K peptide. Incubation of *cyc*⁻ S49 cell membranes with pertussis toxin and ³²P-NAD⁺ resulted in the labelling of a $\sim 39\text{K}$ peptide (lane 3) which was a separate and distinct protein from the N_s subunit labelled by cholera toxin (lane 4).

The above experiment demonstrates that *cyc*⁻ S49 cells contain a pertussis toxin substrate, and that in this respect they resemble other cells possessing an inhibitory regulatory system that affects adenylyl cyclase^{12–17}. To determine whether this substrate is functionally related to inhibition of adenylyl cyclase, *cyc*⁻ S49 cells were incubated in the presence of the toxin for 1–8 h (Fig. 2). In parallel, control cells incubated without toxin were collected at 0, 2, 4 and 8 h after starting the experiment. At each time point, membranes were prepared for adenylyl cyclase assays as described previously¹⁸. Adenylyl cyclase activities in the presence of either Mn or Mg plus forskolin were measured, as well as the ability of GTP and GMP-P(NH)P to inhibit the Mg plus forskolin-stimulated activity. Forskolin was included in the assays in the presence of Mg because it dramatically increases the measurable activity observed, although it is not required for obtaining the inhibitory effects of guanine nucleotides or NaF¹⁸. Figure 2*b* shows that, although the degree of inhibition by GMP-P(NH)P was affected little by toxin treatment, GTP inhibition of *cyc*⁻ S49 cell membrane adenylyl cyclase decreased with time of treatment. By 8 h of incubation with the toxin there was essentially no inhibitory effect of GTP.

Table 1 summarizes the results of seven experiments in which *cyc*⁻ S49 cells were incubated for 8 h in the absence (control) or presence of $0.1 \mu\text{g ml}^{-1}$ pertussis toxin. The actual magnitude

Table 1 Effect of 8 h pertussis toxin treatment of *cyc*⁻ S49 cells on adenylyl cyclase activity and inhibitory responses of cell membranes

	Adenylyl cyclase activity (pmol mg ⁻¹ min ⁻¹)		% Inhibition		
	MnCl ₂	MgCl ₂ + forskolin	GTP	NaF	GMP-P(NH)P
Control	5.3 $\pm$ 1.6*	54.0 $\pm$ 16.1	17.9 $\pm$ 1.7	12.9 $\pm$ 4.8	42.7 $\pm$ 1.7
Toxin-treated	6.4 $\pm$ 2.1	79.0 $\pm$ 29.2	-4.3 $\pm$ 2.2	19.4 $\pm$ 4.4	38.3 $\pm$ 1.2
$\bar{d} \pm \text{s.e.m.}$	-1.1 $\pm$ 0.6	-25.0 $\pm$ 13.5	22.1 $\pm$ 1.1	-6.6 $\pm$ 3.3	4.4 $\pm$ 2.4
<i>t</i> (d.f. = 6)	-1.8	-1.9	20.1	2.0	1.8
<i>P</i>	NS	NS	<0.001	NS	NS

The NaF concentration in the assays was 10 mM; the other conditions of the treatments and assays are described in the text and in Fig. 2 legend. Tests of statistical significance were done using a paired *t*-test. $\bar{d}$, Mean difference between control and toxin-treated values; NS, not significant.

* Values are mean $\pm$ s.e.m. from seven replicate experiments.

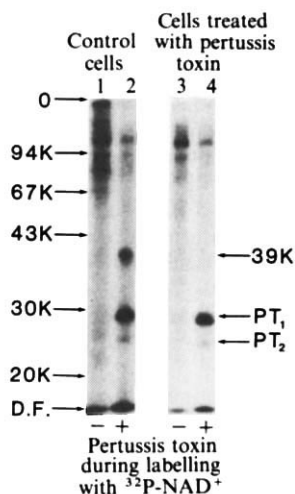


Fig. 3 Effect of pretreatment with pertussis toxin on subsequent labelling with $^{32}\text{P-NAD}^+$ of cyc^- S49 cells in the presence of toxin. cyc^- S49 cells were incubated without (control) or with pertussis toxin for 8 h (as described in Fig. 2 legend) and membranes prepared as described in Fig. 1 legend. The conditions for pertussis toxin treatment of the membranes were the same as in Fig. 1 except that $\sim 20 \mu\text{g}$ of membrane protein and $25 \mu\text{g ml}^{-1}$ toxin were used in the incubation.

of the specific activity of adenylyl cyclase in these separate experiments varied by as much as 10-fold, regardless of whether activities were determined in the presence of MnCl_2 or MgCl_2 plus forskolin. This variability, which is associated with the individual growth of cyc^- cells, seems to be related to the batch of horse serum available and the growth state of the cells at the moment of collection. Specific adenylyl cyclase activities of separate membrane batches derived from a single growth of cells, such as shown in Fig. 2, varied considerably less (never by more than twofold) and did not correlate with any known variable. Despite this variation in specific activity of the adenylyl cyclase in the membranes we prepared, the degree of inhibition by GTP, NaF or GMP-P(NH)P in membranes from control cells was relatively constant (Table 1). Moreover, as in Fig. 2, the results summarized in Table 1 show a consistent effect of pertussis toxin on GTP-mediated inhibition, reducing it to a level not significantly different from 0, with no significant change in the effects of NaF or GMP-P(NH)P.

To determine whether there is a causal relationship between the ADP-ribosylation of the 39K peptide in cyc^- membranes by pertussis toxin and the effect of the toxin on intact cells, the following experiment was performed. cyc^- S49 cell membranes from cells treated for 8 h with pertussis toxin, or control cells incubated without toxin, were incubated with $^{32}\text{P-NAD}^+$ in the presence or absence of pertussis toxin. Figure 3 shows that in control cell membranes there was, once again, a 39K peptide labelled by the toxin (compare lane 2 with lane 1). Membranes from cells previously exposed to pertussis toxin, however, no longer had an ADP-ribosylatable band at M_r 39,000, suggesting that at least one of the effects of the toxin on intact cyc^- S49 cells was to ADP-ribosylate this peptide. In this experiment, which used a higher pertussis toxin concentration and a lower membrane concentration than that in Fig. 1, significant labelling of $\sim 29\text{K}$ peptide, as well as minor labelling of a $\sim 26\text{K}$ peptide, was seen in both toxin-treated and control membranes subsequently incubated with pertussis toxin. These are two of the pertussis toxin subunits (PT_1 and PT_2 in Fig. 3) which are autolabelled in the conditions of the experiment (R.D.S. and C.R.M., unpublished observation).

Thus, the data presented here substantiate the hypothesis that N_i is a regulatory component of adenylyl cyclase which is separate and distinct from N_s . By at least five criteria, cyc^- cells do not contain N_s . Their adenylyl cyclase is stimulated neither by guanine nucleotides nor by NaF^{5,19}, nor is it affected by

cholera toxin treatment^{18,20}; they contain neither radioimmunoassayable 42–45K subunits of N_s (ref. 21) nor cholera toxin substrate for ADP-ribosylation²² and they lack β -adrenergic receptor regulation by guanine nucleotides and Mg (refs 23, 24). The only data which suggest that cyc^- cells might contain an altered form of N_s are the findings that guanine nucleotides and NaF inhibit cyc^- cell membrane adenylyl cyclase activity¹⁸. However, the results of this study correlate the inhibitory effects in the cyc^- system with those seen in other biologically relevant systems where hormones inhibit adenylyl cyclase^{12–17}. In both cases, pertussis toxin treatment leads to attenuation of inhibition and the concurrent ADP-ribosylation of a $\sim 40\text{K}$ peptide. The presence of pertussis toxin substrates in other intact systems (that is, those having a functional N_s) suggests the existence of a N_i (ref. 12) (as do the kinetic arguments for the separate regulation of inhibitory and stimulatory responses^{1,10,11}), but does not constitute proof of this hypothesis as these substrates may represent previously unrecognized subunits of N_s or receptor-related proteins which modulate the activity of N_s . As cyc^- S49 cells do not contain N_s (refs 5, 19–24), the N_i -like activity seen here must be regulating the adenylyl cyclase catalytic unit directly and therefore must represent a transmembrane signalling protein different from N_s . This does not imply that N_i and N_s do not influence each other's ability to affect adenylyl cyclase, or that they may not compete for a similar binding site on the catalytic unit, but only that the two types of regulation of adenylyl cyclase reside in the properties of separate regulatory proteins.

The inhibitory effects seen in cyc^- indicate several very striking similarities between N_i and N_s : both proteins are regulated by NaF as well as by guanine nucleotides¹⁸, and in both cases the nonhydrolysable guanine nucleotide analogues are more active than GTP, the natural effector. For both, hormonal effects correlate with stimulation of a low K_m GTPase activity^{33–35}. In both cases the regulatory proteins are affected by an ADP-ribosylating toxin which causes an increase in the cyclic AMP content of their respective target cells. At the molecular level, the effects of these toxins alter GTP-mediated but not GMP-P(NH)P-mediated regulation. In one important respect, however, the effects of pertussis toxin on N_i differ markedly from those of cholera toxin on N_s . Whereas cholera toxin enhances GTP effects so that they are equivalent to those of GMP-P(NH)P, pertussis toxin reverses the effects of GTP on N_i without altering GMP-P(NH)P effects. Thus there are subtle but important differences in the mechanisms of action of N_i and N_s in their regulation of adenylyl cyclase and their mediation of hormonal effects.

This work was supported in part by NIH grants AM-19318 and AM-27685.

Received 23 November 1982; accepted 15 February 1983.

- Cooper, D. M. F. *FEBS Lett.* **138**, 157–163 (1982).
- Londos, C., Cooper, D. M. F. & Rodbell, M. *Adv. Cyclic Nucleotide Res.* **14**, 163–170 (1981).
- Jakobs, K. H., Aktories, K. & Schultz, G. *Adv. Cyclic Nucleotide Res.* **14**, 173–187 (1981).
- Pfeuffer, T. *J. Biol. Chem.* **252**, 7224–7234 (1977).
- Ross, E. M. & Gilman, A. G. *J. Biol. Chem.* **252**, 6966–6969 (1977).
- Iyengar, R. & Birnbaumer, L. *Proc. natn. Acad. Sci. U.S.A.* **79**, 5179–5188 (1982).
- Northrup, J. K. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 6516–6520 (1980).
- Sternweis, P. C., Northrup, J. K., Smigel, M. D. & Gilman, A. G. *J. Biol. Chem.* **256**, 11517–11526 (1981).
- Hanski, E., Sternweis, P. C., Northrup, J. K., Dromerick, A. W. & Gilman, A. G. *J. Biol. Chem.* **256**, 12911–12919 (1981).
- Motulsky, H. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 4193–4197 (1982).
- Smith, S. K. & Limbird, L. E. *J. Biol. Chem.* **257**, 10471–10478 (1982).
- Katada, T. & Ui, M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3129–3133 (1982); *J. Biol. Chem.* **257**, 7210–7216 (1982).
- Hazeki, O. & Ui, M. *J. Biol. Chem.* **256**, 2856–2862 (1980).
- Katada, T. & Ui, M. *J. Biol. Chem.* **256**, 8310–8317 (1981).
- Cronin, M. J., Myers, G. A., Dabney, L. G. & Hewlett, E. L. *Progr. 64th A. Meet. Endocr. Soc.*, A857 (1982).
- Garcia-Sainz, J. A. *FEBS Lett.* **126**, 306–308 (1981).
- Katada, T. & Ui, M. *J. Biol. Chem.* **255**, 9580–9588 (1980).
- Hildebrandt, J. D., Hanoune, J. & Birnbaumer, L. *J. Biol. Chem.* **257**, 14723–14725 (1982).
- Bourne, H. R., Coffino, P. & Tomkins, G. M. *Science* **187**, 750 (1975).
- Johnson, G. L., Kaslow, H. R. & Bourne, H. R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3113–3117 (1978).
- Ferguson, K. M., Northrup, J. K. & Gilman, A. G. *Fedn Proc.* **41**, 1407 (1982).
- Johnson, G. L., Kaslow, H. R. & Bourne, H. R. *J. Biol. Chem.* **253**, 7120–7123 (1978).
- Bird, S. J. & Maguire, M. E. *J. Biol. Chem.* **253**, 8826–8834 (1978).

24. Ross, E. M., Maguire, M. E., Sturgill, T. W., Biltonen, R. L. & Gilman, A. G. *J. biol. Chem.* **252**, 5761–5775 (1977).
25. Lai, E., Rosen, O. M. & Rubin, C. S. *J. biol. Chem.* **256**, 12866–12874 (1981).
26. Kaslow, H. R., Johnson, G. L., Brothers, V. M. & Bourne, H. R. *J. biol. Chem.* **255**, 3736–3741 (1980).
27. Stadel, J. M., Shorr, R. G. L., Limbird, L. E. & Lefkowitz, R. J. *J. biol. Chem.* **256**, 8718–8723 (1981).
28. Cassel, D. & Pfeuffer, T. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2669–2673 (1978).
29. Gill, D. M. & Meren, R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3050–3054 (1978).
30. Krawietz, W., Werdan, K. K. & Erdmann, E. *Biochem. Pharmac.* **31**, 2463–2469 (1962).
31. Cooper, D. M. F., Jagus, R., Somers, R. L. & Rodbell, M. *Biochem. biophys. Res. Commun.* **101**, 1179–1185 (1981).
32. Malbon, C. G. & Greenberg, M. L. *J. clin. Invest.* **69**, 414–426 (1981).
33. Aktories, K. & Jakobs, K. H. *FEBS Lett.* **130**, 235–238 (1981).
34. Koski, G. & Klee, W. A. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4185–4189 (1981).
35. Cassel, D. & Salinger, Z. *Biochim. biophys. Acta* **452**, 538–551 (1976).
36. Bourne, H. R., Coffino, P. & Tomkins, G. M. *J. cell. Physiol.* **85**, 603–610 (1975).
37. Laemmli, U. *Nature* **227**, 680–685 (1970).
38. Cowell, J. L., Sato, Y., Sato, H., Amderlin, B. & Manclark, C. R. in *Semin. in Infectious Disease Vol. 4* (eds Robbins, J., Hill, J. & Sadoff, J.) 371–379 (Thieme-Stratton, New York, 1982).
39. Irons, L. I. & MacLennan, P. *Biochim. biophys. Acta* **580**, 175–185 (1979).
40. Yajima, M. *et al. J. Biochem., Tokyo* **83**, 295–303 (1978).
41. Meyer, S. L. *Data Analysis for Scientists and Engineers*, 39–41 (Wiley, New York, 1975).

Mode of action of yeast killer toxins: channel formation in lipid bilayer membranes

Bruce L. Kagan*

Department of Physiology and Biophysics,
Albert Einstein College of Medicine, 1300 Morris Park Avenue,
New York, New York 10461, USA

The toxic action of yeast killer proteins seems to involve selective functional damage to the plasma membrane of the sensitive cell. Physiological effects include leakage of K^+ (refs 1, 2), inhibition of active transport of amino acids^{1,3} and acidification of the cell interior^{1,4}. These effects are strikingly similar to the effects of certain bacterial colicins which have been demonstrated previously to form channels in membranes⁵. Proposed mechanisms of action have usually postulated a limited permeability change induced by the toxin in the plasma membrane^{1,4,6,7}. We report here that a killer toxin from the yeast *Pichia kluyveri* forms ion-permeable channels in phospholipid bilayer membranes, and we propose that the *in vitro* electrophysiological properties of these channels account for the morbid effects observed in intoxicated cells. A preliminary account of this work has appeared elsewhere⁸.

Yeast killer toxins are secreted, acidic proteins (molecular weight 10,000–20,000) encoded by double-stranded RNA plasmids (for reviews see refs 9, 10) which are encapsulated in virus-like particles. Killer toxins are only lethal to selected species of yeast, and the killer-producing strain is immune to the effects of its own toxin⁹. Killer toxins bind rapidly, irreversibly and in an energy-independent fashion to a receptor located in the cell wall of sensitive strains⁹. Inhibition of proton-amino acid co-transport and proton extrusion from the cell has been shown to be an early effect of intoxication by the killer toxin of *Saccharomyces cerevisiae*³, and has been linked to an increase in passive H^+ permeability of the intoxicated cell⁴. These results are consistent with the idea that the killer toxin forms an aqueous pore in the plasma membrane of the target cell, thus discharging the electrochemical proton gradient which drives amino acid and K^+ transport¹¹. We used planar phospholipid bilayer membranes to study the effects of yeast killer toxin. Within a few minutes after addition of killer toxin to one side of a membrane, a dramatic increase in the conductance of the membrane was observed (Fig. 1). The toxin-induced conductance is due to the formation of ion-permeable channels (Fig. 2A). Fifteen to thirty minutes after toxin addition, the conductance reached a steady-state value which depended in a linear fashion on the amount of toxin added to the bath (Fig. 3). This

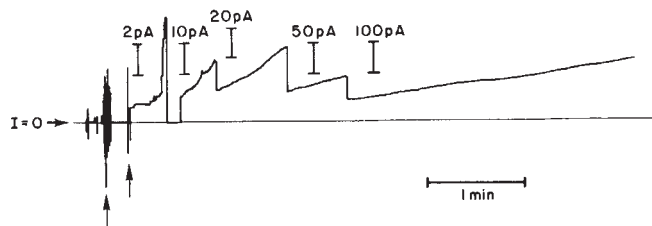


Fig. 1 Current response of lipid membrane to killer toxin. After formation of the asolectin (soybean phospholipid, lecithin type II; Sigma)²⁷ membrane as described previously^{14,28} in 100 mM NaCl, 5 mM dimethylglutaric acid, 1 mM $MgCl_2$, 0.2 mM EDTA, all adjusted to pH 4.65, 20 μ l of partially purified killer toxin were added to the *cis* compartment (vol $\approx$ 4 ml) to a final concentration of 150 ng ml⁻¹ (first arrow). The conductance (reflecting permeability to small ions in the aqueous phase) of a bare membrane is quite low ($\sim$ 10 pS). After addition of killer toxin to the front (or *cis*) compartment to a concentration of 10–10,000 ng ml⁻¹, known voltages were applied across the membrane, and the resulting current responses measured¹⁴. The conductance (G) of the membrane in symmetric salt solutions is defined as current divided by voltage ($G = I/V$), where V is the potential of the *trans* compartment (opposite to the toxin). In the absence of toxin the conductance of the membrane is $\sim$ 10 pS and is indistinguishable from zero on this record. At the second arrow a stimulus of -20 mV (*trans* side negative) was applied and the current instantaneously jumped to a new value indicating the presence of about three killer toxin channels in the membrane. Subsequent stepwise increases of the current (and therefore conductance) were observed as more channels inserted into the membrane. As the current increased there were several reductions in the gain (note calibration bars), and the curve assumed a smoother appearance. The conductance rise shown here is not dependent on voltage, that is, channels can insert without a voltage present as they did between the addition of killer toxin (first arrow) and the application of a voltage (second arrow) where three channels are noted to be present already. Toxin-induced conductance required neither the presence of sterols nor negatively charged lipids in the membrane. Ion selectivity was determined by obtaining the zero-current reversal potential (E , where $I = G(V - E)$) for toxin-treated membranes formed with neutral phospholipids in various salt gradients. In a 2.4-fold activity gradient of KCl (270:100 mM), the reversal potential was 12.4 mV (dilute side positive, ideal selectivity 22.4 mV). In a 2.7-fold activity gradient of NaCl (300:100 mM), the reversal potential was 9.0 mV (dilute side positive, ideal selectivity 25.5 mV). These measurements indicate a preference for K^+ over Cl^- of $\sim$ 3.5:1 and for Na^+ over Cl^- of $\sim$ 2:1.

is consistent with the idea that a single toxin molecule may be responsible for channel formation. Replacement of the aqueous phase bathing the membrane with toxin-free solution had no effect on the toxin-induced conductance, suggesting that channels are irreversibly associated with the membrane. Although the process of channel insertion seems to be independent of voltage, once the channels have entered the membrane they respond to voltage in a characteristic fashion (Fig. 4A). At low voltages ($|V| < 50$ mV) the macroscopic instantaneous and steady-state conductances were identical; at higher voltages ($|V| > 50$ mV) the instantaneous conductance fell within seconds to a lower steady-state value (Fig. 4A). Figure 4B shows the instantaneous and steady-state current-voltage relations for a toxin-treated membrane. Similar voltage-dependent behaviour has been described for channels from bacteria^{12,13} and mitochondria¹⁴.

Unlike the highly selective ion-permeable channels of nerve and muscle¹⁵, the killer toxin channel discriminates rather poorly among common physiological ions. Reversal potential measurements give the following selectivity sequence: $K^+ > Na^+ > Ca^{2+} > Cl^- > 0$. Although we have not directly measured the H^+ permeability, the fact that the cations K^+ , Na^+ and Ca^{2+} can permeate the channel makes it probable that H^+ is permeant also.

The results presented above demonstrate that partially purified extracts of killer toxin from the yeast *P. kluyveri*¹⁶

* Present address: Department of Psychiatry, UCLA Neuropsychiatric Institute, 760 Westwood Plaza, Los Angeles, California 90024, USA.

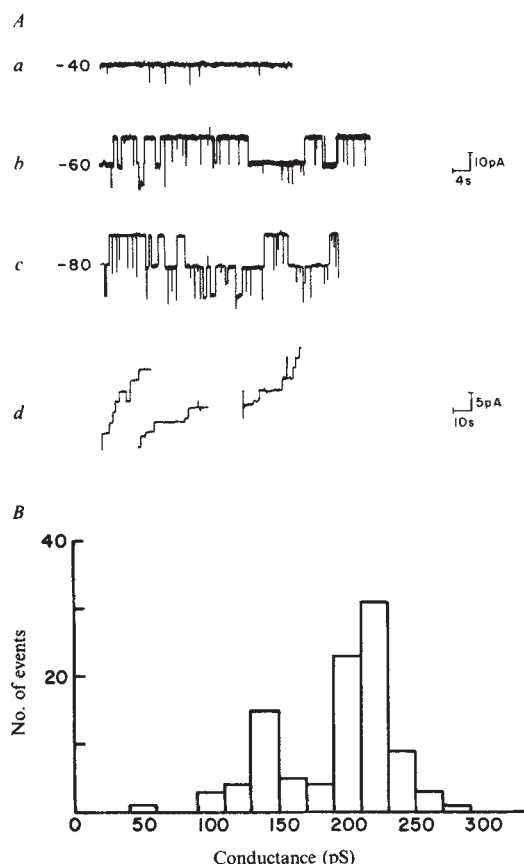


Fig. 2 A, Current (conductance) fluctuations due to the presence of killer toxin channels in the membrane. B, Histogram of the conductance fluctuation sizes observed. At low toxin concentrations, or at short times after toxin addition, there were discrete fluctuations of the steady-state current through the membrane at a given voltage. The amplitude of these current fluctuations was much too large to result from the action of a diffusing ion carrier, and therefore must have been due to the insertion of stable ionic pathways (or channels) in the membrane²⁹. Records a–c are from the same membrane doped with killer toxin at a final concentration of 40 ng ml⁻¹. The aqueous phases contained 100 mM KCl, 5 mM dimethylglutaric acid, 2 mM MgCl₂, 0.1 mM EDTA, all adjusted to pH 4.05. The membrane was formed from a solution of 0.9% asolectin, 0.1% diphytanoyl phosphatidylcholine in hexane. Note that as the absolute magnitude of the voltage was increased, the channels present stayed open a smaller fraction of the time. At higher absolute values of the voltage, the channels were open even less of the time (data not shown). Note that two major jump sizes were observed (see B). d, Single-channel records in an asolectin membrane surrounded by a solution of 100 mM CaCl₂, 5 mM dimethylglutaric acid, 0.1 mM EDTA, all adjusted to pH 4.05. Killer toxin was added to a final concentration of 150 ng ml⁻¹ and the record was obtained 5 min later with the voltage clamped to -10 mV. Note that two jump sizes were observed as in 100 mM KCl. Data in B were from the same membrane as the top three traces of A. Note that there are two main peaks of conductance size centred around 140 pS and 220 pS. Whether these two peaks represent two distinct molecular species or two conductance states of the same channel has not been determined.

induce the formation of ion-permeable channels in lipid bilayer membranes. These channels are due to the active killer toxin itself as similarly prepared extracts from non-killer strains, and toxin preparations inactivated by heat (100 °C for 5 min) or pH (7.4 for 15 min) did not induce channel formation. We suggest that the channel activity we have described can account for the *in vivo* effects of killer toxin.

(1) The killer toxin channel has approximately the right conductance to account for the *in vivo* K⁺ efflux data of Skipper and Bussey¹⁷ and Middelbeek *et al.*¹.

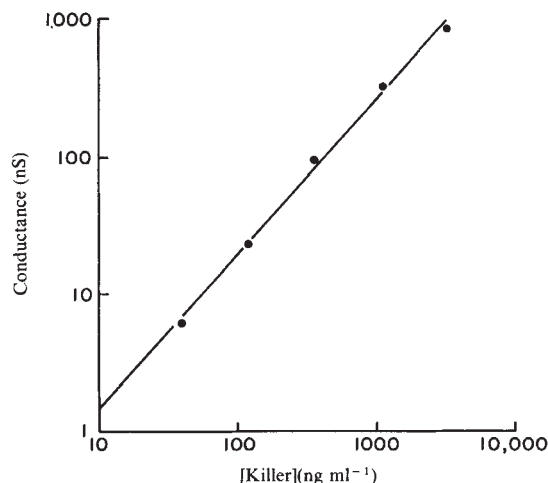


Fig. 3 Double logarithmic plot of steady-state killer toxin-induced conductance against killer toxin concentration in the aqueous phase. The data were obtained on a single asolectin membrane 30 min after each new toxin addition to the concentration indicated. The voltage was held at -10 mV throughout. The aqueous phases were 100 mM KCl, 5 mM Tris, 2 mM MgCl₂, 0.1 mM EDTA, all adjusted to pH 7.0. The slope of the line is 1.13.

(2) The relative lack of ion selectivity we have observed for the killer toxin channel is consistent with the *in vivo* physiology. Macroscopic electroneutrality requires that the observed efflux of K⁺ be accompanied by a corresponding influx of cations (for example, Na⁺) or efflux of anions (such as Cl⁻). As the killer toxin channel is permeable to all these ionic species, one need not invoke other permeability pathways to explain the observed ionic fluxes.

(3) Given the normally low passive permeabilities of the microbial plasma membrane for K⁺, Na⁺ and H⁺ (ref. 18) it is likely that a lethal dose of killer toxin induces a substantial 'leak' pathway for these ions in the plasma membrane. This would naturally lead to a decline in the normal, metabolically maintained, ionic gradients existing across the yeast plasma membrane^{19,20}. The extent and rate of decline of these gradients depend on the size of the toxin-induced leak relative to the proton pumping ability of the cell. Direct *in vivo* measurements of proton pumping⁴ indicate that the toxin-induced leak is large enough to inhibit immediately the net pumping of protons from the cell. This also suggests that the toxin-induced leak is large enough to dissipate any pre-existing membrane potential. We believe that this depolarization of the cell is responsible for the observed immediate inhibition of amino acid transport as the actively transported species is believed to be positively charged, and its transport is strongly dependent on the existence of a (inside negative) membrane potential²¹.

(4) The experiments of Middelbeek *et al.*¹¹ showing that intoxicated cells can be 'rescued' by plating in media containing physiological K⁺ and H⁺ concentrations suggest that the lethal effect of killer toxin is the loss of intracellular K⁺ and the decrease in intracellular pH, after which the cells can no longer be rescued. Similar rescue experiments performed by Kopecky *et al.*²² with *Escherichia coli* demonstrated that loss of intracellular K⁺ and Mg²⁺ were the lethal lesions in intoxication by colicins K and E1 which are also known to be channel formers⁵.

There is a strong resemblance between yeast killer toxins and the bacterial colicins of the E1 functional class. Both are plasmid-encoded toxins lethal to only a limited spectrum of closely related species. In both cases, there is an immunity system (protein?) which protects the toxin-producing cell from its suicidal tendencies. Finally, both classes of toxin induce formation of relatively non-selective ion-permeable channels in the plasma membrane, thus inhibiting vital amino acid transport and changing the intracellular ionic milieu.

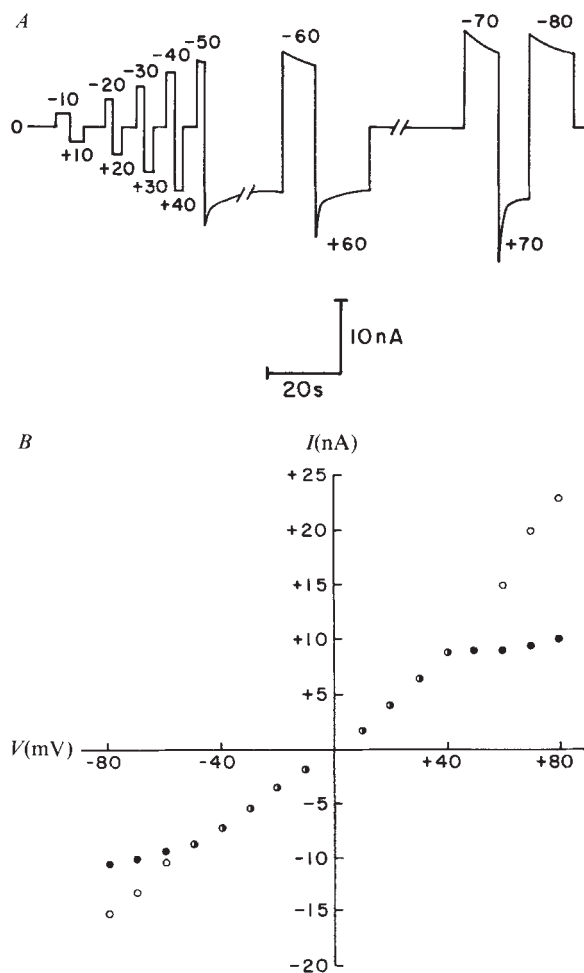


Fig. 4 *A*, Voltage-dependent behaviour of killer toxin-treated membrane. The current responses to different applied voltages are shown. The membrane is the same as that of Fig. 1, observed 15 min later in the steady state. Note that the instantaneous and steady-state responses are identical up to $\sim \pm 40$ mV. At voltages of greater absolute value, the current relaxes with time to a steady-state value smaller than the instantaneous value. This is reflected on the single-channel level by a higher proportion of time spent by a channel in the 'off' state (see Fig. 2*A*). *B*, Instantaneous (○) steady-state (●) *I*-*V* relationship of killer toxin-treated membrane. Data are from the membrane of *A*. Note that there is some rectification of the instantaneous *I*-*V* curve also, with greater currents at (*trans* side) positive voltages than at negative voltages. At voltages of even greater absolute value ($|V| > 120$ mV) more than 90% of the instantaneous conductance can be turned off (data not shown).

Several differences between the two classes of toxins are noteworthy: (1) killer toxins are small, acidic proteins which are secreted by the yeast cell⁹, while the colicins are larger, have varying isoelectric points²³, and are not released until the producer cell lyses²⁴. (2) The killer channels have a single channel conductance which is two orders of magnitude greater than colicin channel conductances²⁵. As a yeast cell has a volume about 1,000 times greater than that of a bacterial cell, the larger conductance of the killer channel allows the lethal leakage of ions to proceed at a reasonable pace. At low multiplicities, killer toxin-induced K^+ efflux takes only ~ 10 times (rather than 1,000 times) longer than colicin-induced K^+ efflux²⁶. (3) Colicin channels exhibit strongly voltage-dependent behaviour in planar lipid membranes⁵, whereas killer channels only appear weakly voltage-dependent (Fig. 4*B*) and the voltage dependence only occurs at relatively high voltages. The role of voltage dependence in the action of both toxins remains obscure.

We thank G. D. Vogels and E. Middelbeek for providing samples of killer toxin and for helpful discussions. We also thank C. L. Slayman for sharing his insights about pH regulation. This is report 1313 from the Galvani Institute for Electrobiology and we thank the director, Dr Alan Finkelstein, for making its resources available to us. The work was supported in part by NIH grant 5T32GM7288.

Received 13 September 1982; accepted 31 January 1983.

- Middelbeek, E. J., Stumm, C. & Vogels, G. D. *Antonie van Leeuwenhoek* **46**, 205-220 (1980).
- Bussey, H. & Skipper, N. *J. Bact.* **124**, 476-483 (1975).
- De la Pena, P., Barros, F., Gascon, S., Ramos, S. & Lazo, P. S. *Biochem. biophys. Res. Commun.* **96**, 544-550 (1980).
- De la Pena, P., Barros, F., Gascon, S., Lazo, P. S. & Ramos, S. *J. biol. Chem.* **256**, 10420-10425 (1981).
- Schein, S. J., Kagan, B. L. & Finkelstein, A. *Nature* **276**, 159-163 (1978).
- Bussey, H. *J. gen. Microbiol.* **82**, 171-179 (1974).
- Kotani, H., Shinmyo, A. & Enatsu, T. *J. Bact.* **129**, 640-650 (1977).
- Kagan, B. L. & Finkelstein, A. *Biophys. J.* **37**, 208a (1982).
- Bussey, H. *Adv. microb. Physiol.* **22**, 93-122 (1981).
- Wickner, R. B. *Plasmid* **2**, 303-322 (1979).
- Middelbeek, E. J., Crutzen, A. Q. H. & Vogels, G. D. *Antimicrob. Ag. Chemother.* **18**, 519-524 (1980).
- Schindler, H. & Rosenbusch, J. P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3751-3755 (1978).
- Ehrenstein, G., Lecar, H. & Nossel, R. *J. gen. Physiol.* **55**, 119-133 (1970).
- Schein, S. J., Colombini, M. & Finkelstein, A. *J. Membrane Biol.* **30**, 99-120 (1976).
- Hille, B. in *Handbk of Physiology* Vol 1, Pt 1 (eds. Brookhart, J. M. & Mountcastle, V. B.) 99-136 (1977).
- Middelbeek, E. J., Hermans, J. M. H. & Stumm, C. *Antonie van Leeuwenhoek* **45**, 437-450 (1979).
- Skipper, N. & Bussey, H. *J. Bact.* **129**, 668-677 (1977).
- Slayman, C. L. *J. gen. Physiol.* **49**, 93-116 (1965).
- Sanders, D. & Slayman, C. L. *J. gen. Physiol.* **80** (in the press).
- Sanders, D., Hanse, U.-P., & Slayman, C. L. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5903-5907 (1981).
- Seaston, A., Carr, G. & Eddy, A. A. *Biochem. J.* **154**, 669-676 (1976).
- Kopecky, A. L., Copeland, C. P. & Lusk, J. E. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4631-4634 (1975).
- Konisky, J. in *The Bacteria* Vol. 6 (eds. Ornston, L. N. & Sokatch, J. R.) 71-136 (Academic, London, 1978).
- Jakes, K. S. & Model, P. *J. Bact.* **138**, 770-778 (1979).
- Kagan, B. L. thesis, Yeshiva Univ., New York (1982).
- Wendt, L. *J. Bact.* **104**, 1236-1241 (1970).
- Kagawa, Y. & Racker, E. *J. biol. Chem.* **246**, 5477-5487 (1971).
- Montal, M. *Meth. Enzym.* **32b**, 545-554 (1974).
- Hladky, S. B. & Haydon, D. A. *Biochim. biophys. Acta* **274**, 294-312 (1972).

Surface immunoglobulin-negative B-cell precursors detected by formation of antibody-secreting colonies in agar

Christopher J. Paige

Basel Institute for Immunology, Grenzacherstrasse 487, CH-4005 Basel, Switzerland

The development of semi-solid *in vitro* cloning assays has helped distinguish the different stages in early haematopoietic differentiation. The progenitors of erythrocytes, granulocytes, macrophages and megakaryocytes have been quantitated and characterized in such systems¹⁻⁶ but until now, similar assays for progenitors of antibody-producing B lymphocytes have not been established despite many reports describing the properties of B-cell precursors which proliferate and differentiate either *in vivo* in adoptive hosts or in *in vitro* liquid culture systems⁷⁻²⁵. A semi-solid agar assay is described here which permits a murine B-cell precursor to develop into a colony containing antibody-secreting cells after 7-11 days in culture. The precursor cells were found in fetal liver and could be clearly distinguished from mature clonable B cells. This assay thus provides a method to quantitate functional B-cell precursors and establish the requirements for the generation of B lymphocytes.

Approximately 20-25% of mature surface immunoglobulin positive (sIg⁺) B lymphocytes placed in semi-solid agar and stimulated by mitogens will proliferate to form recognizable colonies^{26,27}. Of these clones, 25-50% contain IgM-secreting cells by day 4 of culture^{26,27}. During ontogeny colony-forming B cells first appear on days 16-17 of gestation at the time sIg⁺

B lymphocytes^{23,28} emerge. Fetal liver cells obtained from younger fetuses do not form agar colonies in these conditions. However, when such cultures were established in agar over an underlying adherent layer of 14-day fetal liver cells, clones of proliferating cells were detected (Table 1). Most of these colonies appeared to be myeloid clones containing granulocytes and macrophages. To test the possibility that colonies containing antibody-secreting cells were also present, the ability of these clones to secrete antibody was analysed using the agar plaque assay previously described in detail^{26,27}. This method, in which the agar colonies are overlaid with protein A-coupled sheep red blood cells, developing antisera and complement, allows the detection of colonies containing cells belonging to the B lineage even if these clones are infrequent^{26,27}. These experiments revealed that a small percentage of the adherent-layer dependent colonies contained IgM-secreting cells (Table 1).

Previous experiments with adult cells have shown that lipopolysaccharide (LPS) significantly augments both the number of B cells which clone in agar and the ability of the cells in these colonies to secrete antibody^{26,27,29}. To test when LPS was effective in the fetal liver system, the addition of LPS to the cultures was delayed. Table 1 shows that a delay of 3 days had little effect on the number of plaque forming clones detected. Thus LPS significantly augments the ability of fetal liver colonies to generate plaque forming cells, but it may only be effective after B cells have arisen.

Table 1 Proliferation and plaque formation by 14-day fetal liver cells in response to various stimuli

Culture conditions	Total colonies and clusters (per 10^5 14-day fetal liver cells)	Plaque forming colonies
1. LPS	6 ± 2	<1
2. LPS+adherent layer cells	65 ± 12	6 ± 2
3. Adherent layer cells, no LPS	55 ± 9	<1
Adherent layer cells	63 ± 5	4 ± 1
+LPS added after 6 days		
Adherent layer cells	86 ± 8	9 ± 1
+LPS added after 3 days		
Adherent layer cells	93 ± 9	10 ± 4
+LPS from beginning		

The first row of data are a summary of five experiments in which the response of 10^5 or 2×10^5 14-day fetal liver cells to LPS ($25 \mu\text{g ml}^{-1}$) was measured after 8 days in double layer culture using methods already reported²⁶ based on the B-cell colony assay described previously^{31,32}. In these experiments Iscove's modified Dulbecco's medium (IMDM) was prepared daily from Dulbecco's modified Eagle's medium and supplemented with 0.5% delipidated and deionized bovine serum albumin (BSA, Behring), $5 \mu\text{g ml}^{-1}$ transferrin and 7.5% fetal calf serum³³. In each experiment 5–20 replicate cultures were examined. Colonies contained >20 cells and clusters contained >4 cells. (C57BL/6 $\times$ DBA/2) F_1 embryos were used in these experiments and the method used to obtain fetal liver cells was as described previously^{23,24}. The morning of vaginal plug detection was considered day 0. The second row of data is a summary of nine experiments in which the response of 14-day fetal liver cells to LPS and adherent cells was measured after 8 days in double layer agar culture. The adherent layer was established by culturing 2×10^6 14-day fetal liver cells in 1 ml of IMDM supplemented as described above (except with 0.05% BSA) for 5 h at 37°C and 5% CO_2 . Titration experiments revealed 2×10^6 cells to be optimum although the actual number of cells which adhere has not been determined. After incubation the non-adherent cells were removed by suction and double layer agar cultures were set up on top of the adherent layer. Occasional colonies and plaques could be detected emerging from the adherent layer (or remaining non-adherent cells) but their number was insignificant compared with the colonies and plaques generated from the added top layer cells. The remaining data are from a single experiment designed to determine when LPS was effective in the cultures. LPS (0.1 ml ; final concentration of $25 \mu\text{g ml}^{-1}$) was added at the times indicated. Replicate cultures (6–10) were examined after a total of 8 days in culture.

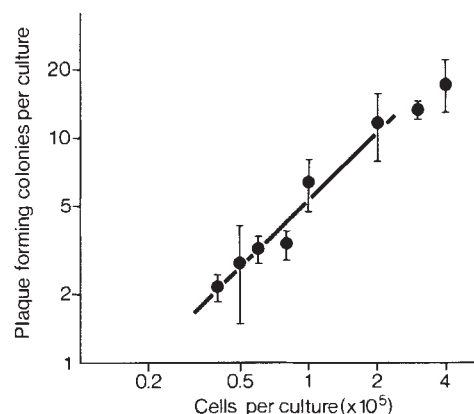


Fig. 1 Relationship between the number of 8-day plaque forming colonies and the number of 14-day fetal liver cells plated. Each point represents the pooled data from individual experiments (total of 6) in which 5–20 replicate cultures were examined. The least squares regression line generated from the logarithms of the arithmetic means of the points from 4×10^4 to 2×10^5 is shown. Its slope is 0.99, consistent with plaque forming colony dependence on a single progenitor cell.

To test whether the efficiency of detection of plaque forming colonies was dependent on cell concentration, cell numbers were titrated in six experiments (Fig. 1). The slope of the log: log relationship between the number of input cells and the number of plaque forming colonies suggested that only one cell type was limiting in the conditions described. Occasional plaque forming colonies were detected without adherent layers but only at very high seeding densities ($\geq 3 \times 10^5$ per ml). The frequency of B-cell progenitors detected by this assay is 1 in 18,000 14-day fetal liver cells. Similar experiments with 15-day fetal liver yielded an estimate of 1 in 6,000. Plaque forming colonies were also generated from 12- and 13-day fetal liver cells; however, the higher ratio of myeloid colonies to plaque forming colonies in the younger fetal liver populations resulted in deteriorating culture conditions when high cell numbers were seeded, making frequency estimates less reliable.

Several experiments were conducted to determine the time course of plaque formation as well as the size and surface properties of the cell which initiates colony formation. Figure 2 shows that fetal liver cells do not generate plaque forming colonies until day 6 or 7 of agar culture whereas plaque forming colonies derived from adult spleen arise on day 2 of culture and peak on day 4. In control experiments, the time course of plaque formation of spleen cells was not changed by the presence of a fetal liver adherent layer. Sedimentation velocity analysis of the fetal liver cells which gave rise to plaque forming colonies indicated they were heterogeneous with a modal rate of sedimentation of 5 mm per h, similar to that of large pre-B cells detected in other assay systems^{7,21,24,25} (Fig. 3). In contrast, the modal rate of sedimentation of clonable B cells from adult bone marrow, measured either by colony counts or by plaque formation, was 2.9 mm per h, in agreement with previous reports^{7,21,24,25} (data not shown).

The surface characteristics of the cell which initiates colony formation were examined by depleting the fetal liver cells of cells binding to plastic Petri dishes coated first with mouse anti-rat immunoglobulin and then selected rat anti-mouse monoclonal antibodies following the panning procedure of Kincaid *et al.*³⁰. In four experiments in which fetal liver cells were exposed to rat anti-mouse monoclonal antibody 14.8, the number of plaque forming colonies was reduced by an average of 61% compared with populations exposed to the mouse anti-rat antibody alone. Monoclonal antibody 14.8 has previously been shown to deplete progenitors which give rise to B cells in liquid culture or in adoptive hosts³⁰. Similar experiments with rat anti-mouse monoclonal antibody 33-60, which is specific for μ

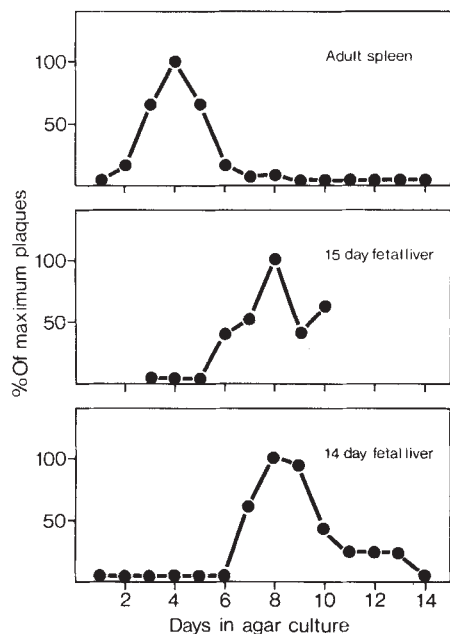


Fig. 2 Time course of plaque formation by colonies derived from fetal liver and adult spleen cells. In all cases LPS ($25 \mu\text{g ml}^{-1}$) was included in the cultures. Each point represents the mean of 4–10 replicate plates. Maximum responses were $4.5 \text{ per } 10^5$ for 14-day fetal liver; $29 \text{ per } 10^5$ for 15-day fetal liver; and $17.5 \text{ per } 10^3$ for adult spleen.

heavy chains, showed no reduction in plaque forming colonies. In contrast, exposure of adult spleen cells to monoclonal 33-60 resulted in a decrease of $>95\%$ in B-cell colonies. These data indicate that the majority of cells which give rise to 8-day plaque forming colonies are μ negative but do express the antigen detected by antibody 14.8.

Morphological analysis of the colonies which contained plaque forming cells was carried out on day 8 of culture using methods of fixation and staining previously described²⁶. The plaque forming colonies were composed of cells which were indistinguishable from lymphocytes found in B-cell colonies derived from adult spleen.

Attempts to obtain a conditioned medium from the fetal liver adherent cells which could overcome the adherent-layer dependence of this assay have been unsuccessful. However, conditioned medium obtained by culturing 12- or 13-day mouse placentas in medium with or without serum for 48–72 h was active in supporting plaque forming colonies. Crude supernatant 0.5% was sufficient to support >200 myeloid colonies and 7 plaque forming colonies per 2×10^5 14-day fetal liver cells cultured. In these cases, however, cell numbers of 2×10^5 per ml or greater were required to obtain consistent results. Since these supernatants also contained factors which stimulated myeloid progenitors, it is possible that their action on B-cell progenitors was indirect.

A functional assay is described here which allows the detection, quantitation and clonal analysis of B-cell precursors. That the colony forming cell in this assay is a B-cell progenitor and not a rare B cell is suggested by the delayed onset of antibody secretion, delayed reactivity to LPS, large size as assessed by sedimentation rate, and the failure to adhere to anti- μ antibodies. The estimated frequency of this cell type is a minimum value although it is higher than estimates using liquid culture assays¹⁷. In the mature B-cell colony assay, only 1 in 4 sIg⁺ cells proliferates, and of these 50% at most, can generate an IgM plaque at any particular time^{26,27}. If these characteristics are already programmed in B-cell precursors, then the estimated frequency of 1 in 18,000 14-day fetal liver cells may be low.

The finding that the number of plaque forming colonies is directly proportional to the number of input cells is consistent

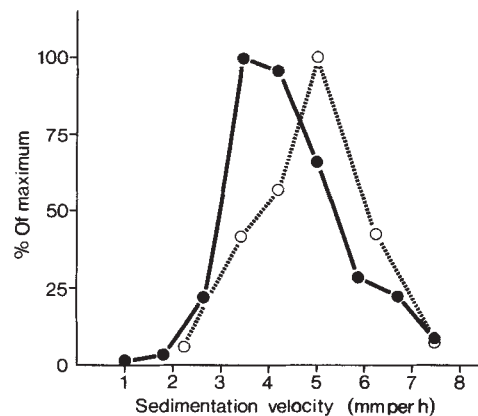


Fig. 3 Sedimentation velocity analysis of 15-day fetal liver cells showing profiles of total nucleated cells (●—●) and 8-day plaque forming colonies (○—○)^{24,34}. The data are based on 6–10 replicate cultures. Peak plaque forming colony response was 38 plaque forming colonies per 5×10^4 cells cultured. The total number of plaque forming colonies was calculated based on incidence values and the total number of cells per fraction and is expressed as a % of the maximum value. In this experiment 50% of the input number of plaque forming colony cells were recovered.

with the view that the progenitor cell is not dependent on other colony types which arise concomitantly. Whether the fetal liver adherent cells or the placental cells provide a physiological microenvironment for B-cell differentiation remains unknown. However, the ability to quantitate functional B-cell precursors and observe their clonal differentiation into mature antibody-secreting B cells will help elucidate the cellular stages and developmental events of B lymphocytes.

I thank H. Skarvall for technical assistance; C. Brügger for preparing the manuscript; R. Grützmann for providing the cell line 33-60; P. Kincade for providing monoclonal antibody 14.8; E. Collins for monitoring the mouse mating colony; R. Gisler, W. Haas, and F. Melchers for comments and N. Iscove for many helpful suggestions. The Basel Institute for Immunology was founded and is supported by F. Hoffman-La Roche and Co.

Received 13 December 1982; accepted 11 February 1983.

- Bradley, T. R. & Metcalf, D. *Aust. J. exp. Biol. med. Sci.* **44**, 287–299 (1966).
- Pluznik, D. H. & Sachs, L. *Expl. Cell Res.* **43**, 533–540 (1966).
- Stephenson, J. R., Axelrad, A. A., McLeod, D. L. & Shreeve, M. M. *Proc. natn. Acad. Sci. U.S.A.* **68**, 1542–1546 (1971).
- Axelrad, A. A., McLeod, D. L., Shreeve, M. M. & Heath, D. A. in *Hemopoiesis in Culture* (ed. Robinson, W.) 226–234 (US Government, Washington DC, 1974).
- Iscove, N. N., Sieber, F. & Winterhalter, K. H. *J. cell. Physiol.* **83**, 309–320 (1974).
- Metcalf, D., MacDonald, H. R., Odartchenko, N. & Sordat, B. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1744–1748 (1975).
- Lafleur, L., Miller, R. G. & Phillips, R. A. *J. exp. Med.* **135**, 1363–1374 (1972).
- Lafleur, L., Miller, R. G. & Phillips, R. A. *J. exp. Med.* **137**, 954–966 (1973).
- Nossal, G. V. & Pike, B. L. *Immunology* **25**, 33–45 (1973).
- Stutman, O. in *Microenvironmental Aspects of Immunity* (eds Jankovic, B. D. & Isakov, K.) 19–26 (Plenum, New York, 1973).
- Osmond, D. G. & Nossal, G. J. V. *Cell. Immun.* **13**, 132–145 (1974).
- Owen, J. J. T., Cooper, M. D. & Raff, M. C. *Nature* **249**, 361–363 (1974).
- Gelfand, M. C., Asofsky, R. & Paul, W. E. *Cell. Immun.* **14**, 460–469 (1974).
- Ryser, J. E. & Vassalli, P. *J. Immun.* **113**, 719–728 (1974).
- Hämmerling, U., Chin, A. F. & Abbott, J. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2008–2012 (1976).
- Rozing, J., Brons, N. H. C. & Benner, R. *Cell. Immun.* **29**, 37–53 (1977).
- Melchers, P. *Eur. J. Immun.* **7**, 482–486 (1977).
- Kincade, P. W., Moore, M. A. S., Lee, G. & Paige, C. J. *Cell. Immun.* **40**, 294–302 (1978).
- Burrows, P. D., Kearney, J. F., Lawton, A. R. & Cooper, M. D. *J. Immun.* **120**, 1526–1531 (1978).
- Scheid, M. P., Goldstein, G. & Boyse, E. A. *J. exp. Med.* **147**, 1727–1743 (1978).
- Lau, C. Y., Melchers, P., Miller, R. G. & Phillips, R. A. *J. Immun.* **122**, 1273–1277 (1979).
- Teale, J. M. & Mandel, T. E. *J. exp. Med.* **151**, 429–445 (1980).
- Paige, C. J., Kincade, P. W., Moore, M. A. S. & Lee, G. *J. exp. Med.* **150**, 548–563 (1979).
- Paige, C. J., Kincade, P. W., Shinefeld, L. A. & Sato, V. L. *J. exp. Med.* **153**, 154–165 (1981).
- Kincade, P. W., Lee, G., Paige, C. J. & Scheid, M. P. *J. Immun.* **127**, 255–260 (1981).
- Paige, C. J. & Skarvall, J. *immun. Meth.* **52**, 51–61 (1982).
- Paige, C. J. in *B and T Cell Tumors* (ed. Vitetta, E.) 179–183 (Academic, New York, 1982).
- Johnson, G. R., Metcalf, D. & Wilson, J. W. *Immunology* **30**, 907–914 (1976).
- Kincade, P. W., Ralph, P. & Moore, M. A. S. *J. exp. Med.* **143**, 1265–1270 (1976).
- Kincade, P. W., Lee, G., Watanabe, T., Sun, L. & Scheid, M. P. *J. Immun.* **127**, 2262–2268 (1981).
- Metcalf, D. *et al. J. exp. Med.* **142**, 1534–1549 (1975).
- Kincade, P. W. in *Immunological Techniques Applied to Aging Research* (eds Adler, W. H. & Nordin, A. A.) 85–97 (CRC, West Palm Beach, 1981).
- Iscove, N. N. & Melchers, F. *J. exp. Med.* **147**, 923–933 (1978).
- Miller, R. G. & Phillips, R. A. *J. cell. Physiol.* **73**, 191–199 (1969).

A secreted phosphoprotein marker for neoplastic transformation of both epithelial and fibroblastic cells

Donald R. Senger*, Bonnie B. Asch†, Barbara D. Smith‡, Carole A. Perruzzi* & Harold F. Dvorak*

* Department of Pathology, Beth Israel Hospital and Harvard Medical School, and the Charles A. Dana Research Institute, Beth Israel Hospital, Boston, Massachusetts 02215, USA

† Department of Experimental Pathology, Roswell Park Memorial Institute, Buffalo, New York 14263, USA

‡ Veterans Administration Outpatient Clinic and Boston University School of Medicine, 17 Court St, Boston, Massachusetts 02108, USA

A wide variety of virally and spontaneously transformed fibroblasts secrete a major transformation-related phosphoprotein with a molecular weight (MW), depending on the species of origin, of about 62,000 (62K)¹⁻³. Markedly elevated extracellular levels of this major ³²P-labelled protein are not simply linked to exponential growth¹ but instead are associated directly with transformation. The phosphoprotein is not antigenically related to p60^{src}, p60^{c-src} or simian virus 40 (SV40) non-viral T antigen², and it is further distinguishable from SV40 non-viral T antigen (pp 53) on the basis of its electrophoretic mobility. In this study we have compared a variety of normal and transformed epithelial cells for secretion of this ³²P-labelled protein and have found that this marker distinguishes neoplastic from preneoplastic and normal mouse mammary epithelium and also identifies highly tumorigenic cells derived from guinea pig bile duct epithelium and rat liver epithelium. Because the classical phenotypic properties commonly associated with transformation of fibroblasts cannot be generally used to discriminate tumorigenic from non-tumorigenic epithelial cells⁴⁻⁹, this phosphoprotein, which identifies tumorigenic cells of both fibroblastic and epithelial origin, is likely to be of particular importance.

We first analysed normal, preneoplastic and neoplastic cells derived from mouse mammary epithelium for secretion of the phosphoprotein. Primary cell cultures were prepared from normal, preneoplastic and neoplastic BALB/c mouse mammary tissue¹⁰. Normal mammary glands were taken from primiparous mice at 15–18 days of gestation. Preneoplastic tissue was taken from D2 hyperplastic alveolar nodules (D2 HAN). The D2 HAN line originally arose in a hormonally stimulated mouse¹¹ and has been maintained by serial transplantation in mammary fat pads of syngeneic BALB/c mice. Approximately 50% of the D2 HAN implants give rise to mammary adenocarcinomas¹¹ and these primary neoplasms (D2 tumours) served as the source of tumour tissue. When grown in the mammary fat pad all of these cultured cells produce the same type of tissue from which they were originally derived^{10,12}. Primary cultures were labelled with ³²P-orthophosphate, and the culture media were analysed for the presence of the 62K MW transformation-related secreted phosphoprotein. Figure 1A shows that the primary mammary D2 tumour cells release markedly elevated levels of a 63K MW ³²P-labelled protein compared with normal and preneoplastic (D2 HAN) cells. Furthermore, this phosphoprotein is antigenically related to the 62K MW transformation-related secreted phosphoprotein secreted by cells of fibroblastic origin.

We next examined two previously established guinea pig bile duct carcinoma lines for secretion of the phosphoprotein. Line 1 and line 10 tumours^{15,16} originated in syngeneic strain 2 guinea pigs which had been given the water-soluble carcinogen diethylnitrosamine in their drinking water. Ascites variants¹⁵ of the two tumour lines were used here. The line 1 cells are weakly

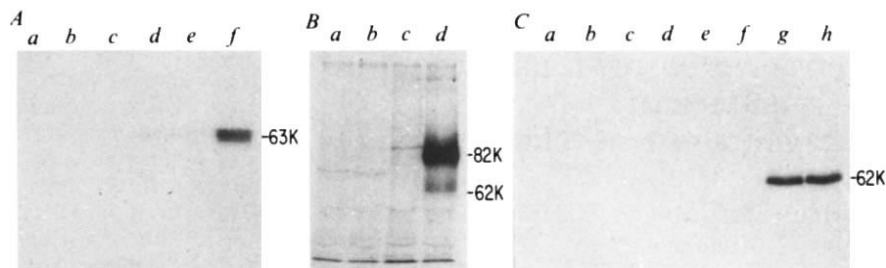
tumorigenic in that tumour nodules grow and regress 10–14 days after intradermal implantation in strain 2 guinea pigs^{15,17}. In contrast, intradermal injection of line 10 cells invariably results in progressive tumour growth, metastasis, and death of the animal¹⁶⁻¹⁸. Figure 1B shows that only the highly malignant line 10 cells secrete 62K and 82K phosphoproteins that cross-react with specific antiserum³ raised against the 62K MW phosphoprotein secreted by transformed fibroblasts. The precise relationship between the 82K and 62K phosphoproteins is unknown, but the weakly tumorigenic line 1 cells do not release detectable levels of either phosphoprotein.

We have also studied a series of rat liver cell lines with a broad range of tumorigenicity. The K-16 line was derived from normal Sprague-Dawley rat liver epithelium¹⁹, and these cells are not tumorigenic in newborn²⁰ or irradiated weanling Sprague-Dawley rats (B.D.S., unpublished results). The W-8 line was derived by treating K-16 cells with the carcinogen *N*-acetoxy-2-acetylaminofluorene²⁰. In contrast to K-16 cells which are not tumorigenic, W-8 cells were found to be tumorigenic with low incidence (see Fig. 1 legend). W-8 cells were recently passaged serially in irradiated weanling Sprague-Dawley rats and highly tumorigenic derivatives were obtained. M-3 and M-4 lines derived from these serial passages form tumours with 83% and 100% incidence respectively, in contrast to the W-8 line which forms tumours with 33% incidence (B.D.S., unpublished). We examined K-16, W-8, M-3 and M-4 cells for secretion of the transformation-related phosphoprotein and found that only the highly tumorigenic M-3 and M-4 lines secrete substantial levels of a 62K MW ³²P-labelled protein (Fig. 1C). Longer autoradiographic exposures (twice that of Fig. 1C) indicated that the less tumorigenic W-8 cell population does secrete low levels of the ³²P-labelled protein whereas no secretion of the phosphoprotein by the normal K-16 cells was detected. Note that the mechanism behind the observed differences between normal and tumour cells might involve any of the following: (1) tumorigenic cells synthesize and secrete more 62K phosphoprotein than their normal counterparts, (2) tumorigenic and normal cells synthesize and secrete the 62K phosphoprotein equally, but tumorigenic cells secrete the protein in a more highly phosphorylated form (10-fold or more), or (3) tumorigenic and normal cells synthesize the same amount of phosphoprotein and phosphorylate it equally, but tumorigenic cells secrete the phosphoprotein preferentially. All of these possibilities are of interest, and experiments are in progress to distinguish among them.

We have used the method of Cooper and Hunter²¹ to determine that all of the phosphate associated with the transformation-related secreted phosphoproteins is esterified through alkali-labile bonds (1M KOH, 55 °C, 2 h, data not shown). Phosphoproteins secreted by rat M-3 and M-4 carcinomas, the HSV-NIL8 hamster sarcoma³, and the B77 rat 1 sarcoma³ were analysed. The alkali-lability of the phosphate esters associated with these phosphoprotein bands indicates that phosphoserine may be present. Recently Otto *et al.* have shown by electrophoretic analysis of acid hydrolysates that the 62K phosphoprotein secreted by Abelson virus-transformed NIH 3T3 cells contains only phosphoserine²². It therefore seems unlikely that the phosphoprotein is phosphorylated by tyrosine-specific kinases such as those associated with transformation by various retroviruses²³⁻²⁵.

Finally, the 62–63K MW tumour-related phosphoproteins secreted by the neoplastically transformed mouse mammary tumour cells (D2 tumour) and the highly tumorigenic cells derived from rat liver epithelium (M-3, M-4) are by far the major phosphoproteins secreted by these cells (Fig. 2). Similarly, the analogous transformation-dependent secreted phosphoproteins of transformed fibroblasts are by far their major secreted phosphoproteins (Fig. 2A *c* and ref. 1). It is noteworthy that the 62–63K MW phosphoproteins secreted by D2 tumour and polyoma-transformed BALB/c 3T3 cells display microheterogeneity (Fig. 2A) and that the average mobility of these phosphoproteins differs slightly between transformed

Fig. 1 Immunoprecipitation of ^{32}P -labelled culture medium; autoradiograms of 7.5% Laemmli¹³ polyacrylamide gels. Cells were labelled in serum-free medium for 3 h with ^{32}P -orthophosphate as previously described¹. Volumes of labelled culture medium used for immunoprecipitations were normalized for total incorporation (trichloroacetic acid-precipitable d.p.m.) in the cell layer. Differences were twofold or less. Immunoprecipitations were carried out with previously characterized antiserum³ specific for the 62K MW phosphoprotein secreted by avian sarcoma virus-transformed rat 1 fibroblasts; fixed *Staphylococcus aureus*¹⁴ was used to absorb immune complexes and resulting pellets were washed three times with phosphate-buffered saline. Molecular weights are indicated (markers were bovine serum albumin, 67,000; catalase, 60,000; and ovalbumin, 43,000), and all samples were reduced with dithiothreitol (0.1 M) and heated (95 °C) before electrophoresis. **A:** a-c, preimmune serum; d-f, immune serum. a and d, normal mouse mammary epithelium; b and e, mammary hyperplastic aveolar nodule (D2HAN); c and f, D2 mammary tumour. **B:** a, preimmune serum, line 1 (weakly tumorigenic guinea pig bile duct carcinoma); b, immune serum, line 1; c, preimmune serum, line 10 (highly tumorigenic guinea pig bile duct carcinoma); d, immune serum, line 10. **C:** a-d, preimmune serum; e-h, immune serum. a and e = K-16 (normal rat liver epithelial cells), b and f = W-8 (weakly tumorigenic, derived from K-16), c and g = M-3 (highly tumorigenic, derived from W-8), d and h = M-4 (highly tumorigenic, derived from M-3). In irradiated (700 rad, 3 min) weanling Sprague-Dawley rats, K-16 cells (1×10^6 , s.c.) did not give rise to tumours (0/12 animals); incidence of tumours with the other cell lines were as follows: W-8, 33% (4/12 animals); M-3, 83% (10/12 animals); M-4, 100% (10/10 animals); (B.D.S., unpublished data). Autoradiographic exposure times: panels A and B, 2 days; panel C, 1 day.



fibroblastic cells (Fig. 2A c) and mammary tumour cells (Fig. 2A b), both derived from the same strain of mouse.

In conclusion, elevated extracellular levels of the major ^{32}P -labelled protein is a readily identified characteristic which distinguishes both neoplastically transformed epithelial cells (mouse, guinea pig, rat) and transformed fibroblasts (mouse, rat, hamster, rabbit; see ref. 3) from their normal or non-tumorigenic counterparts. Moreover, this characteristic correlates with transformation which has been induced by a variety of tumour viruses including polyoma, hamster sarcoma virus, avian sarcoma virus, Moloney sarcoma virus², Abelson murine leukaemia virus²², as well as by chemical carcinogens and hormonal stimulation (present work). We do not yet know the function of the phosphoprotein, but the close correlation between this marker and neoplastic transformation of a wide variety of rodent and rabbit cells by a wide variety of agents suggests that it may have a very basic association with tumorigenic potential. Of particular significance is the fact that this marker can be used to distinguish tumorigenic from non-tumorigenic mouse mammary epithelium and yet mouse mammary tumour cells do not display any of the classical phenotypic properties commonly associated with transformation. For example, both normal and neoplastically transformed mouse mammary epithelial cells in primary cultures are morphologically identical²⁶⁻²⁸, form identical polarized monolayers with specialized cell-cell junctional complexes²⁷, have similar growth rates and saturation densities²⁶, and cannot be distinguished on the basis of anchorage-independent growth^{4,29}. The two cell types also lack significant differences in several surface-related features^{10,27,28} and their distributions of microtubules and microfilaments are indistinguishable^{12,28}. Because primary mouse mammary tumour cells display only minimal features of transformation and yet they do secrete markedly increased

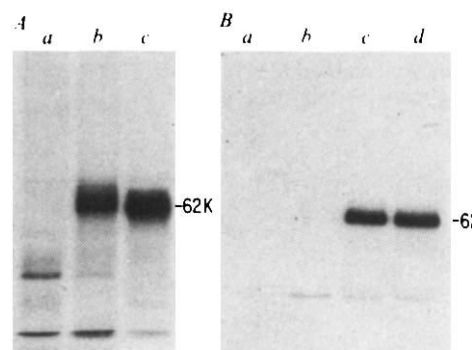


Fig. 2 Total culture medium from cells labelled with ^{32}P -orthophosphate; autoradiograms of 7.5% polyacrylamide gels. Labelling of cells, electrophoresis, and normalization of volumes were as in Fig. 1. **A:** a, BALB/c mouse mammary D2HAN (preneoplastic); b, BALB/c mouse mammary D2 tumour; c, polyoma-transformed BALB/c 3T3 mouse fibroblasts. **B:** a, K-16 (normal rat liver epithelial cells); b, W-8 (weakly tumorigenic cells derived from K-16); c, M-3; d, M-4. M-3 and M-4 are highly tumorigenic lines derived from W-8 (see Fig. 1 and text).

levels of the ^{32}P -labelled protein, it appears that either elevated secretion or phosphorylation of this phosphoprotein is one of the minimal phenotypic alterations required for establishment of tumorigenic potential.

We thank Dr Daniel Medina for providing the mice used in these studies, Ann Mahoney for technical assistance and Dr B. Zbar for line 1 and line 10 cells. This work was supported by NIH grants CA 28471 to H.F.D., CA 31755 to B.B.A., CA 23540 to B.D.S. and a Veterans Administration merit approved programme (B.D.S.).

Received 10 September 1982; accepted 16 February 1983.

- Senger, D. R., Wirth, D. F. & Hynes, R. O. *Cell* **16**, 885-893 (1979).
- Senger, D. R., Wirth, D. F., Bryant, C. & Hynes, R. O. *Cold Spring Harb. Symp. quant. Biol.* **44**, 651-657 (1980).
- Senger, D. R., Wirth, D. F. & Hynes, R. O. *Nature* **286**, 619-621 (1980).
- Butel, J. S., Dudley, J. P. & Noonan, C. P. in *Cell Biology of Breast Cancer* (eds McGrath, C., Brennan, M. S. & Rich, M. A.) 317-345 (Academic, New York, 1980).
- Hosick, H. L., Anderson, L. W., Angello, J. C. & Danielson, K. G. in *Cell Biology of Breast Cancer* (eds McGrath, C., Brennan, M. S. & Rich, M. A.) 393-412 (Academic, New York, 1980).
- Miller, F. R., Dexter, D. L. & Heppner, G. H. in *Cell Biology of Breast Cancer* (eds McGrath, C., Brennan, M. S. & Rich, M. A.) 413-422 (Academic, New York, 1980).
- Marshall, C. J., Franks, L. M. & Carbonell, A. W. *J. natn. Cancer Inst.* **58**, 1743-1751 (1977).
- San, R. H. C. *et al. Cancer Res.* **39**, 1026-1034 (1979).
- Yuspa, S. H., Hahley-Nelson, P., Koehler, B. & Stanley, J. R. *Cancer Res.* **40**, 4694-4703 (1980).
- Asch, B. B. & Medina, D. *J. natn. Cancer Inst.* **61**, 1423-1430 (1978).
- Medina, D. & DeOme, K. B. *J. natn. Cancer Inst.* **45**, 353-363 (1970).
- Asch, B. B., Medina, D. & Brinkley, B. R. *Cancer Res.* **39**, 893-907 (1979).
- Laemmli, U. K. *Nature* **227**, 680-685 (1970).
- Kessler, S. W. *J. Immun.* **115**, 1617-1624 (1975).

- Zbar, B., Wepsic, H. T., Rapp, H. J., Whang-Peng, J. & Borsos, T. *J. natn. Cancer Inst.* **43**, 821-831 (1969).
- Zbar, B., Bernstein, I., Tanaka, T. & Rapp, H. J. *Science* **170**, 1217-1218 (1970).
- Dvorak, H. F., Dvorak, A. M., Manseau, E. J., Wiberg, L. & Churchill, W. H. *J. natn. Cancer Inst.* **62**, 1459-1472 (1979).
- Hunter, J. T., Ashley, M. P., Rapp, H. J., Sukumar, S. & Zbar, B. *Cancer Immun. Immunother.* **11**, 39-44 (1981).
- Weinstein, I. B., Orenstein, J. M., Gebart, R., Kaighn, M. E. & Stadler, U. C. *Cancer Res.* **35**, 253-263 (1975).
- Yamaguchi, N. & Weinstein, I. B. *Proc. natn. Acad. Sci. U.S.A.* **72**, 214-218 (1975).
- Cooper, J. A. & Hunter, T. *Molec. Cell. Biol.* **1**, 165-178 (1981).
- Otto, G., Baltimore, D. & Hynes, R. O. *J. Virol.* (submitted).
- Hunter, T. & Sefton, B. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1311-1315 (1980).
- Witte, O. N., Dasgupta, A. & Baltimore, D. *Nature* **283**, 826-831 (1980).
- Feldman, R. A., Hanafusa, T. & Hanafusa, H. *Cell* **22**, 757-766 (1980).
- Das, N. K., Hosick, H. L. & Nandi, S. *J. natn. Cancer Inst.* **52**, 849-861 (1974).
- Pickett, P. B., Pielka, D. R., Hamamoto, S. T. & Misfeldt, D. S. *J. Cell Biol.* **66**, 316-332 (1975).
- Asch, B. B., Medina, D., Kretzer, F., Connolly, J. L. & Brinkley, B. R. *Cancer Res.* **40**, 2383-2389 (1980).
- McGrath, C. M., Maloney, T. & Soule, H. D. in *Cell Biology of Breast Cancer* (eds McGrath, C., Brennan, M. S. & Rich, M. A.) 373-391 (Academic, New York, 1980).

Hamster preproglucagon contains the sequence of glucagon and two related peptides

**Graeme I. Bell, Robert F. Santerre*
& Guy T. Mullenbach**

Chiron Corporation, 4560 Horton Street, Emeryville,
California 94608, USA and *Division of Molecular and Cell Biology,
Lilly Research Laboratories, Indianapolis, Indiana 46285, USA

Glucagon is a 29-amino acid polypeptide hormone synthesized by the A cells of the endocrine pancreas¹⁻⁴. Its primary site of action is the liver where it stimulates glycogenolysis, gluconeogenesis and ketogenesis. In mammals, biosynthetic studies have shown that glucagon is derived from a precursor of molecular weight (M_r) approximately 18,000 which is five to six times larger than glucagon⁵. Glucagon-containing polypeptides and immunoreactants of various sizes have also been described from stomach, intestine, brain and salivary gland³. Here, we have determined the structure of hamster pancreatic preproglucagon from the sequence of its cDNA. This 180-

amino acid precursor contains the sequence of glucagon and two glucagon-like polypeptides arranged in tandem. The precursor also contains the sequences of several non-pancreatic glucagon-containing polypeptides which suggests that, in mammals, both pancreatic and non-pancreatic glucagon and glucagon-containing polypeptides may be derived from a common precursor by tissue-specific processing. We have tentatively identified each of the glucagon-like immunoreactants which have been described with respect to the sequence of proglucagon and have proposed a scheme for the processing of pancreatic proglucagon.

The structure of preproglucagon mRNA was deduced from the 1,118 base pair (bp) sequence of the insert in colony pshglu1 (Fig. 1). The sequence contains a single open reading frame beginning at the methionine codon at nucleotides 104–106 (the only one in any of the three frames) which predicts the sequence of the 180-amino acid preproglucagon. Thus, the coding region of hamster preproglucagon mRNA is 540 nucleotides. The 3'-untranslated region of the mRNA is 475 bases and contains two polyadenylation signals, AAUAAA^{6,7}, nucleotides 811–816 and 1,098–1,103, beginning 408 and 21 bases, respectively, before the poly(A) tract. The 5'-untranslated region is at least 103 nucleotides. Electrophoresis of glyoxylated islet RNA in an agarose gel⁸ and subsequent hybridization⁹ indicated that hamster preproglucagon mRNA is 1,250 bases (data not shown). As the cloned cDNA insert is 1,118 nucleotides exclu-

Fig. 1 Primary structure of hamster pancreatic preproglucagon mRNA and protein and comparison with anglerfish preproglucagon and porcine GLI-1. The predicted amino acid sequence of hamster preproglucagon is numbered by designating the first amino acid of hamster GLI-1 as 1. The amino acids constituting the signal peptide are given negative numbers. The basic dipeptides which may be involved in processing are boxed. The regions corresponding to glucagon and GLP-1 and -2 are indicated. The number of the nucleotide at the end of each line is indicated. The two AAUAAA sequences in the 3'-untranslated region are underlined. Hamster and anglerfish¹⁴ preproglucagon are aligned and colons indicate gaps

[illegible]

introduced to maximize homology. The amino acid differences between hamster and porcine GLI-1¹⁰ (proglucagon(1-69)) are indicated.

Methods: RNA was prepared²³ from islets of Langerhans obtained from the pancreases of 400 female Syrian hamsters (*Mesocricetus auratus*)²⁴. The tissue preparation was 50% islets and contained $\sim 2 \times 10^7$ cells. The yield of poly(A)-containing RNA was 35 μ g. Double-stranded cDNA was prepared as described by Land *et al.*²⁵ and inserted into the *Pst*I site of pBR322 using the GC-tailing technique²⁶. Transformation of *Escherichia coli* strain HB101 generated a library of 25,000 tetracycline-resistant transformants; 1,000 of these were grown in arrays on Whatman 541 filter paper²⁷. Following amplification of the plasmid DNA *in situ*, the colonies were screened for those containing glucagon sequences by hybridization²⁸ with a ³²P-labelled heptadecadeoxynucleotide²⁹ pool whose sequences 3' GT(T/C) ACC (G/A)A(A/G/C/T) TAC TA(A/G) TG 5' were complementary to all 32 possible sequences of mRNA encoding glucagon (24–29). Two colonies containing plasmids with inserts of 1,200 and 1,000 bp hybridized: pshglu1 and 2, respectively. The colonies were rescreened with the insert in pshglu1 and no additional hybridizing colonies were observed. The frequency of colonies containing plasmids coding for hamster insulin and somatostatin was also determined by hybridization with human insulin³⁰ and somatostatin¹⁵ probes: 20/500 and 2/1,000 colonies hybridized with the insulin and somatostatin probes, respectively. The sequence of the insert in pshglu1 was determined by the procedure of Maxam and Gilbert³¹ on both strands except for 165 bp and 210 bp at the 5' and 3' ends, respectively, and across all restriction sites used to initiate sequence determinations. A partial sequence of the insert in pshglu2 indicated no differences in the region from nucleotides 240 to 329.

his ser gln gly thr phe thr ser asp Glucagon
his asp gln phe gln arg his ala gln gly thr phe thr ser asp GLP-1
his ala asp gly val phe thr ser asp PH-27
tyr ala asp ala ile phe thr asn ser pGRF
tyr ser lys tyr leu asp ser arg ala gln asp phe val gln Glucagon
val ser thr leu asp ser leu ala thr arg asp phe ile asn GLP-2
leu ser arg leu arg asp ser ala arg leu gln arg leu gln Secretin
thr thr arg leu arg lys ile arg gln asp phe val asp VIP
trp ser ile ala met asp lys ile arg gln asp phe val asp GIP
phe arg leu leu gln leu ser ala lys tyr leu gln PH-27
tyr arg lys val leu gln leu ser ala arg lys leu gln pGRF

trp leu met asn thr trp leu val lys gln thr lys ile thr asp lys lys
gln leu val ser ile leu val trp leu leu ala gln lys gln lys ser asp trp lys his asn
GIP PH-27 pGRF
asp ile met ser arg gln gln gly gln ser asn gln arg gly pGRF

ile thr gln ala arg ala leu pGRF

Fig. 2 Comparison of the sequences of hamster glucagon-like polypeptides with other members of the glucagon-secretin family. The peptides are: human glucagon (all mammalian glucagons seem to have identical sequences)³²; hamster GLP-1 and -2 (this paper); porcine secretin, vasoactive intestinal polypeptide (VIP), gastric inhibitory polypeptide (GIP) and PH-32-34, and human pGRF³⁵.

ive of the poly(A) tract, it may represent a nearly full-length copy of the mRNA. Also, as an approximately 900-base form of hamster pancreatic preproglucagon mRNA was not observed, the proximal AAUAAA, nucleotides 811-816, is not a normal signal for polyadenylation.

The organization of hamster preproglucagon was determined by comparing its sequence with that of other glucagon-containing polypeptides. This analysis suggested that the first 20, mainly hydrophobic, amino acids constitute the signal peptide. Thus, hamster preproglucagon is 160 amino acids and its predicted molecular weight of 18,675 is in good agreement with the value of 18,000 determined by Patzelt *et al.*⁵ for the rat precursor. Glucagon is proglucagon(33-61) and is flanked by 52 and 99 amino acids at its amino and carboxy terminus, respectively. Proglucagon(1-69) possesses 90% homology with porcine intestinal glucagon-like immunoreactant 1 (GLI-1) or

glucagon¹⁰ and is probably the corresponding hamster protein (Fig. 1). Proglucagon(1-30) is 80% homologous to a porcine polypeptide, called gliadin-related pancreatic polypeptide (GRPP), which is secreted from the pancreas concomitantly with glucagon¹¹. Proglucagon(33-69) is the pancreatic glucagon precursor initially described by Tager and Steinert¹² and which Battalio *et al.*¹³ recently characterized from porcine intestine. Because pancreatic proglucagon contains the sequences of both pancreatic and intestinal glucagon-containing polypeptides, a common precursor may be synthesized in both tissues. The carboxy-terminal segment of proglucagon, residues 70-160, contains two glucagon-like peptides (GLP) of 37 and 35 amino acids, GLP-1 and -2 (Fig. 1). Each polypeptide is flanked by a pair of basic amino acids which can be sites of proteolytic processing. However, there is no evidence to suggest that the Arg-Arg at residues 109, 110 and 124, 125 are cleaved. In fact, the Arg-Arg at residues 49, 50 in the glucagon moiety is not cleaved. In addition, spacer oligopeptides of 6 and 13 residues separate glucagon and GLP-1, and GLP-1 and GLP-2, respectively. GLP-1 and -2 are related but not identical to other members of the glucagon-secretin family of gastro-intestinal hormones which have been described (Fig. 2). Lund *et al.*¹⁴ have characterized an anglerfish pancreatic preproglucagon. This 124-amino acid precursor (M_r 14,500) is 56 amino acids smaller than hamster preproglucagon and this difference is due to the absence of the 13-amino acid spacer peptide and second glucagon-like peptide (GLP-2) in the anglerfish precursor (Fig. 1). Interestingly, this is the first example in which the organization of a pro-hormone has not been conserved during vertebrate evolution (compare mammalian and fish preproinsulin and preprosomatostatin^{15,16}). Also, in contrast to mammals, anglerfish has another preproglucagon of ~12,500 M_r (refs 17, 18); however, its sequence has not been reported. Thus, there may be at least three different types of pancreatic proglucagon in vertebrates. Comparing hamster and anglerfish preproglucagon, the signal peptide, the amino-terminal peptide (corresponding to GRPP), glucagon and GLP-1 possess 25, 10, 69 and 48% amino acid homology and 47, 33, 76 and 66% nucleotide homology, respectively. The low level of sequence conservation in the signal peptide region is not unexpected because the absolute sequence of this region is not as important as the maintenance of its hydrophobic character¹⁹. Although the sequence of the amino-terminal peptide, that is, proglucagon(1-30), is not conserved, the size is and this segment may be required for proper processing of the precursor. Interestingly, in mammals, the

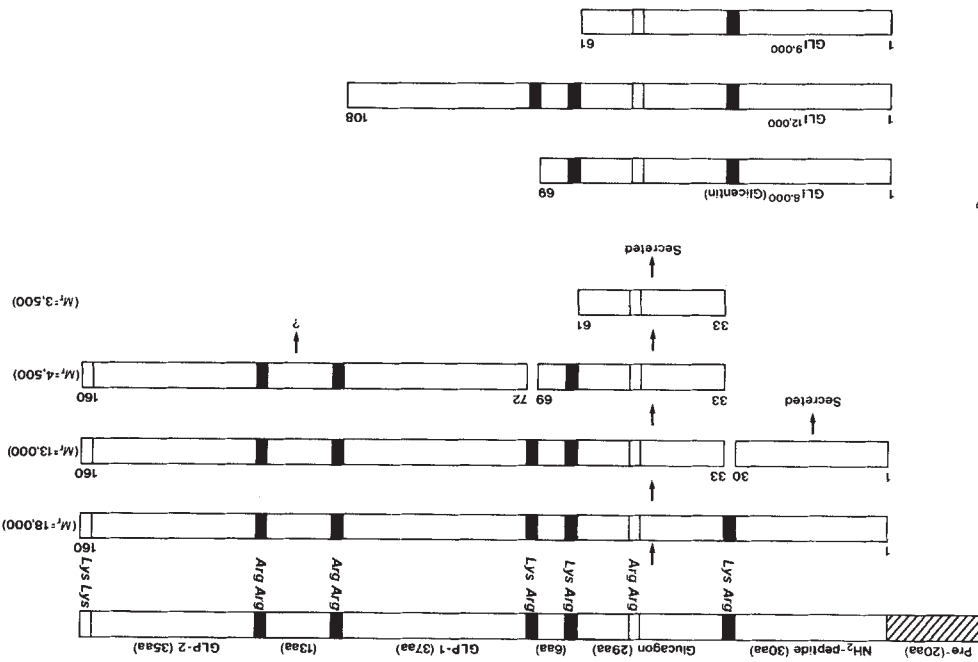


Fig. 3 Schematic representation of the processing of pancreatic pre-glucagon and the structure of glucagon-containing polypeptides. *a*, Possible pathway for the proteolytic processing of pancreatic proglucagon. The basic dipeptides are indicated and those which are potential sites for cleavage are indicated by arrows. The numbers in parentheses at the end of each line are the sizes of the glucagon-containing intermediates determined by Patzelt *et al.*⁵. The numbers above the lines are the amino acids at the ends of the polypeptide in relation to the sequence of preproglucagon (Fig. 1). *b*, Structure of non-pancreatic glucagon-containing polypeptides. GLI 9,000 and GLI 12,000 are major polypeptides in the intestine and GLI 9,000 accumulates in the serum of animals with renal failure^{3,20}.

sequence of proglucagon(1–30) is conserved to a greater extent than the C-peptide of proinsulin. For example, there is 80% homology between hamster and porcine proglucagon(1–30) and only 48% homology between their insulin C-peptides. The corresponding values in a comparison of hamster and human are 83% and 65%, respectively (G.I.B., unpublished). The spacer peptide which separates glucagon and GLP-1 is 6 amino acids in hamster and 5 in anglerfish and the sequence is different. The GLP-1 region possesses extensive homology between hamster and anglerfish, especially the segment corresponding to hamster proglucagon(78–100). However, hamster GLP-1 has a 6-amino acid amino-terminal extension which is absent in anglerfish. The sequence conservation in the GLP-1 region suggests that this peptide has a biological function. We have also compared hamster GLP-2 with the anglerfish GLP and they possess only 29% amino acid sequence homology. The 5'- and 3'-untranslated portions of hamster and anglerfish mRNA possess no significant regions of homology.

The pancreas and intestine are the major sites of synthesis of glucagon and glucagon-containing polypeptides. Figure 3 is a possible scheme for the processing of proglucagon in the pancreas. This model is based on the structure of preproglucagon presented here, the sizes and order of appearance of intermediates in the processing of rat proglucagon and the assumption that processing occurs at basic dipeptides. As indicated, both glucagon and proglucagon(1–30) are secreted¹¹. However, the fate of proglucagon(72–160) and the two glucagon-like polypeptides is unknown. Intestinal glucagon-containing polypeptides are present at less than 1% the levels of pancreatic glucagon and the major polypeptides have M_r s of 8,000 and 12,000 (ref. 20). The 8,000- M_r protein is GLI-1 and corresponds to proglucagon(1–69) (Figs 1, 3). The sequence of the 12,000- M_r polypeptide has not been determined but its size, biochemical and immunological properties^{3,20} are consistent with it being proglucagon(1–108), and thus it would include both glucagon and GLP-1 (Figs 1, 3). In addition, a 9,000- M_r peptide with glucagon-like immunoreactivity has been described which accumulates in the plasma of animals with renal failure^{1–3,20}. Its size, biochemical and immunological properties suggest that it may correspond to proglucagon(1–61) (Fig. 3).

Glucagon and insulin have a major role in the regulation of plasma glucose levels. The biological role(s) of the intestinal glucagon-containing polypeptides is unclear although GLI-1 seems to inhibit gastric acid secretion²¹. Both the pancreatic and intestinal glucagon-containing polypeptides are probably derived from a common precursor which is processed differently in these two tissues. The processing of proglucagon can potentially generate at least 11 unique polypeptides, 8 of which have been identified biochemically or immunochemically. As the sequence of proglucagon is now known, it will be possible to synthesize polypeptides and to produce antisera to specific segments or polypeptides contained within the precursor²². The processing of the precursor in different tissues and the function of this family of glucagon-containing polypeptides in normal and disease states can then be critically examined. The difference in structure between mammalian and anglerfish proglucagon is unusual. The presence of an additional glucagon-like peptide in the mammalian precursor suggests that duplication or loss of a segment of the proglucagon gene has occurred. Examination of the structure of the gene may elucidate the evolutionary history of this hormone.

We thank our colleagues at Lilly Research Laboratories who helped in the isolation of the islets, Ms M. Quiroga and N. Fong for their technical assistance, Dr W. J. Rutter for his encouragement and support, Drs Leslie B. Rall and Pablo Valenzuela for their assistance in writing and Dr R. Najarian and Ms Maureen Appling who prepared the figures and typed the manuscript.

Received 2 December 1982; accepted 16 February 1983.

I. Foa, P., Bajaj, J. & Foa, N. (eds) *Glucagon: Its Role in Physiology and Clinical Medicine* (Springer, Berlin, 1977).

2. Unger, R. H. & Orci, L. (eds) *Glucagon. Physiology, Pathophysiology and Morphology of the Pancreatic A-Cell* (Elsevier, New York, 1981).
3. Conlon, J. M. *Diabetologia* **18**, 85–88 (1980).
4. Unger, R. H. & Orci, L. *New Engl. J. Med.* **304**, 1518–1524 (1981).
5. Patzelt, C., Tager, H. S., Carroll, R. S. & Steiner, D. F. *Nature* **282**, 260–266 (1979).
6. Proudfoot, N. J. & Brownlee, G. G. *Nature* **263**, 211–214 (1976).
7. Fitzgerald, M. & Shenk, T. *Cell* **24**, 251–260 (1981).
8. Carmichael, G. C. & McMaster, G. K. *Meth. Enzym.* **65**, 380–391 (1980).
9. Thomas, P. S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201–5205 (1980).
10. Thim, L. & Moody, A. J. *Regulatory Peptides* **2**, 139–150 (1981).
11. Moody, A. J., Holst, J. J., Thim, L. & Lindkaer, J. *Nature* **289**, 514–516 (1981).
12. Tager, H. S. & Steiner, D. F. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2321–2325 (1973).
13. Bataille, D. et al. *FEBS Lett.* **146**, 79–86 (1982).
14. Lund, P. K., Goodman, R. H., Dee, P. C. & Habener, J. F. *Proc. natn. Acad. Sci. U.S.A.* **79**, 345–349 (1982).
15. Shen, L. P., Pictet, R. L. & Rutter, W. J. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4575–4579 (1982).
16. Chan, S. J. et al. *J. biol. Chem.* **256**, 7595–7602 (1981).
17. Lund, P. K., Goodman, R. H. & Habener, J. F. *J. biol. Chem.* **256**, 6515–6518 (1981).
18. Shields, D., Warren, T. G., Roth, S. E. & Brenner, M. J. *Nature* **289**, 511–514 (1981).
19. Steiner, D. F., Quinn, P. S., Chan, S. J., Marsh, J. & Tager, H. S. *Ann. N.Y. Acad. Sci.* **343**, 1–16 (1980).
20. Tager, H. S. & Markese, J. J. *J. biol. Chem.* **254**, 2229–2233 (1979).
21. Kirkegaard, P. et al. *Nature* **297**, 156–157 (1982).
22. Lerner, R. A. *Nature* **299**, 592–596 (1982).
23. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294–5299 (1979).
24. Santerre, R. F. et al. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4339–4343 (1981).
25. Land, H., Grez, M., Hauser, H., Lindenmaier, W. & Schutz, G. *Nucleic Acids Res.* **9**, 2251–2266 (1981).
26. Villa-Komaroff, L. et al. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3727–3731 (1978).
27. Gergen, J. P., Stern, R. H. & Wensink, P. C. *Nucleic Acids Res.* **7**, 2115–2136 (1979).
28. Comb, M., Seeburg, P. H., Adelman, J., Eiden, L. & Herbert, E. *Nature* **295**, 663–666 (1982).
29. Beauchage, S. L. & Caruthers, M. H. *Tetrahedron Lett.* **22**, 1859–1862 (1981).
30. Bell, G. I. et al. *Nature* **282**, 525–527 (1979).
31. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499–560 (1980).
32. Dayhoff, M. O. *Atlas of Protein Sequence and Structure* Vol. 5, Suppl. 2, 125–126 (National Biomedical Research Foundation, Washington, DC, 1976).
33. Tatemoto, K. & Mutt, V. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6603–6607 (1981).
34. Jornval, H. et al. *FEBS Lett.* **123**, 205–216 (1981).
35. Guillemin, R. et al. *Science* **218**, 585–587 (1982).

A new troponin T and cDNA clones for 13 different muscle proteins, found by shotgun sequencing

Scott D. Putney*, Walter C. Herlihy† & Paul Schimmel*

* Department of Biology and † Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Complete amino acid sequences have been established for 19 muscle-related proteins and these proteins are each sufficiently abundant to suggest that their mRNA levels are about 0.4% or higher. Based on these considerations, a simple theoretical analysis shows that clones for most of these proteins can be identified within a complementary DNA library by sequencing cDNA inserts from 150–200 randomly selected clones. This procedure should not only rigorously identify specific clones, but it could also uncover amino acid sequence variants of major muscle proteins such as the troponins^{1–6}. We have determined sequences for about 20,000 nucleotides within 178 randomly selected clones of a rabbit muscle cDNA library, and report here that in addition to finding sequences encoding the two known skeletal muscle isoforms of troponin C^{7–9}, we have discovered sequences encoding two forms of troponin T. Over the region of nucleotide sequence overlap in the troponin T clones, the new isotype diverges significantly from its counterpart¹⁰. Altogether, clones for 13 of the 19 known muscle-specific proteins were identified, in addition to the clone for the new troponin T isotype.

To identify a clone for a particular protein by DNA sequencing, the nucleotide sequence of a cDNA clone encoding a portion of the protein sequence must be determined. We determined sequences of cDNA fragments isolated from a library of rabbit muscle cDNA cloned into M13 phage¹¹. Before cloning, the cDNA was restricted with *MspI*, *TaqI* or *Sau3AI* so that cDNA fragments of ~250 base pairs (bp) were actually cloned. Sequences of ~110 nucleotides from 178 different phage inserts were determined. The sequences were translated

by computer into the six possible polypeptide sequences (three for the sequenced strand and three for its complement) and compared with complete sequences of 19 proteins prominent in rabbit muscle. These proteins are the 13 whose cDNAs were identified (see below) plus rabbit myoglobin¹², LC-1 and LC-3 myosin light chains^{13,14}, parvalbumin^{15,16}, adenylate kinase (S. Kuby, personal communication) and the M-isozyme of lactate dehydrogenase from pig muscle¹⁷. In addition, partial sequences of the M-isozyme of creatine kinase (S. Kuby, personal communication and W.C.H. and K. Biemann, unpublished observations) and of the sarcoplasmic Ca^{2+} -pump ATPase¹⁸⁻²⁰ were checked.

Identification of clones was made on the basis of a match between the translated amino acid sequence and the protein sequence. Matches are of sufficient length to assure positive identification. For example, the probability of a chance occurrence of a specific sequence encoding 12 amino acids in our compilation of 20,000 nucleotides is $\sim 10^{-5}$. All matches are at least this length and most are several times longer.

Table 1 lists the proteins whose sequences match sequences encoded by one or more clones, and shows the number of nucleotides sequenced and the location in the protein sequence of the encoded amino acids. For most of the proteins, more than one clone was found, with actin by far the most common. On the other hand, only one clone was found for glyceraldehyde 3-phosphate dehydrogenase, myosin heavy chain, triose phosphate isomerase and α - and β -tropomyosin. In the case of rabbit glyceraldehyde 3-phosphate dehydrogenase, the clone was found using the pig protein sequence and there is 95% homology in the region sequenced between the proteins from these two animals.

Note that, in Table 1, the sequences obtained are usually scattered about the various coding regions, with no particular bias towards sequences encoding the amino- or carboxy-termini. This implies that all of the recutting and recloning operations generated a good randomization so that fragments from various parts of the respective coding regions were generated.

The nucleotide and amino acid sequences of the regions sequenced are given in Table 2 and form the basis for an unequivocal identification of each clone. Although only a portion of each of the cDNAs was cloned, we have found that the fragments in Table 2 are each of sufficient length to screen (by hybridization) for full-length cDNAs.

The frequency of detecting a clone for a protein by shotgun sequencing should roughly approximate the weight per cent of its messenger RNA in the total poly(A) RNA from which the cDNA is synthesized. In addition to mRNA levels, both the number of restriction sites in the cDNA and the length of the 3' untranslated region presumably affect the per cent of each clone in the library.

For cases where they have been measured, mRNA levels correlate with the frequency of detection of the corresponding cDNA clone. Of poly(A) RNA, actin mRNA constitutes 8% (ref. 21), LC-1 and LC-2 myosin light chain mRNAs constitute 2.0% (ref. 22) and myosin heavy chain mRNA constitutes 1.1% (ref. 23) in chick skeletal muscle, chick embryonic heart tissue and rat myotubes, respectively. The frequency of detection of actin cDNA, LC-2 myosin cDNA and myosin heavy chain cDNA was 18.5% (that is, 33 out of 178 clones), 1.7% and 0.56%, respectively. Although our data base is small, these frequencies of detection are within a factor of about 2 of the mRNA levels, suggesting that the frequencies of detection correlate with mRNA levels.

Troponin is a component of the thin filament which confers Ca^{2+} sensitivity on actomyosin in the presence of tropomyosin²⁴. It is a complex of three subunits (troponin C, I and T) and several cDNA clones were detected for each.

For the region encoding amino acids 132-146 of troponin C, two cDNA sequences differing at 8 of 43 consecutive nucleotides were found (Table 3). Five of these differences are not evident in the amino acid sequence, but the others suggest a

heterogeneity at position 133 because one sequence encodes a glutamic acid residue and the other a serine. Troponin C from fast skeletal muscle of rabbit contains a Ser at this position⁷, while that from slow muscle contains a Glu⁹. This suggests that one clone is for fast skeletal muscle troponin C and the other is for slow muscle.

As for troponin T, presumably there are polymorphic forms of this protein²⁻⁶. In addition to identification of the only previously characterized species¹⁰, the sequence data show the existence of a new isotype. There is a 60-nucleotide stretch that decodes into an amino acid sequence which aligns with residues 217-236. Of these 20 amino acids, there are four positions which differ in the two isotypes (Table 3) and these differences are all conservative in nature. At points of overlap in the two cloned nucleotide sequences, there is substantial divergence between the two forms (Table 3). With the new isotype reported here, there are now two troponin T isotypes demonstrated at the sequence level. These two isotypes could be analogous to the fast and slow muscle isotypes of troponin C.

Although previous approaches have been directed at cloning individual muscle protein cDNAs (for example, refs 25-28), the sequencing strategy used here allowed, in one sweep, selection of cDNA clones for the majority of the characterized

Table 1 cDNA clones for specific muscle proteins

Protein	No. of clones obtained	No. of nucleotides sequenced	Location of sequence in protein
α -Actin (375)	5	150	72-120
	1	89	102-130
	3	50	130-145
	8	52	192-208
	3	124	208-247
	4	176	224-280
	2	64	308-328
	4	40	328-340
	2	52	340-356
	1	61	356-374
Aldolase A (361)	2	58	37-55
	2	50	347-361
Glyceraldehyde 3-phosphate dehydrogenase (332)	1	142	32-78
Myosin heavy chain (1,880)	1	81	1,643-1,669
LC-2 myosin light chain (169)	1	103	112-145
	2	38	26-37
Phosphorylase (841)	1	53	657-673
	3	67	721-741
	1	61	735-754
	1	103	769-802
Triose phosphate isomerase (248)	1	84	56-83
α -Tropomyosin (284)	1	132	103-146
β -Tropomyosin (284)	1	130	131-172
Troponin C (159) fast	2	94	73-102
	2	55	132-148
(161) slow	1	49	134-148
Troponin I (178)	2	75	1-24
	1	43	162-174
Troponin T (259)	2	148	49-97
	1	43	97-110
	1	99	233-259
New isotype	2	64	217-236

cDNA clones of 14 different rabbit muscle proteins identified by DNA sequencing. The numbers in parentheses indicate the number of amino acids in the protein. Total RNA was isolated from the back and hind legs of a 1-month-old female rabbit, and the poly(A) fraction purified by oligo(dT) chromatography³¹. Using 15 μg of poly(A) RNA, 15 μg double-stranded cDNA was synthesized with reverse transcriptase (a gift from James Beard) followed by DNA polymerase³². This was treated with S₁ nuclease³³ and fractionated by electrophoresis through a 5% polyacrylamide gel³⁴. DNA longer than 450 bp was eluted³⁵ and 0.4 μg was digested separately with *TaqI*, *MspI* or *Sau3AI*. The *TaqI* and *MspI* fragments were ligated, separately, with 40 ng of M13 vector mp8 (ref. 11) linearized with *AccI*, and the *Sau3AI* was ligated with 40 ng mp8 linearized with *BamHI*. When used to transform host JM103 (ref. 36) each mixture gave ~ 150 phages containing inserts. Single-stranded DNA was prepared from 60, 56 and 62 separate phages from the *TaqI*, *MspI* and *Sau3AI* libraries, respectively. DNA sequences were obtained by the chain termination technique³⁷.

Table 2 Nucleotide and translated amino acid sequences of specific cDNA clones

α -Actin	TCGAGCACGGCATTACCAACTGGGACACATGGAGAAGATCTGGACCAACCTCTTACAACGAGCTGGCGTGGCCCCGAGGACACCCACTCTGCTCACCAGGCCCCCTCAACCCCAAGGCCAACCGCGAG GluHisGlyIleIleThrAsnTrpAspAspMetGluLysIleTrpHisIleThrPheTyrAsnGluLeuArgValAlaProGluGluHisProThrLeuLeuThrGluAlaProLeuAsnProLysAlaAsnArgGlu 72 AAGATGACCAATGATCATGTTTCGAGACCTTCAACGTGGCCCATGTACGTGGCCATCCAGGCCGTGCTGTCCTGTACGCCCTCCGG LysMetThrGlnIleMetPheGluThrPheAsnValProAlaMetTyrValAlaIleGlnAlaValLeuSerLeuTyrAlaSer 145 GATCCTCACTGAGCGTGGCTACTTCTCGTACCACAGCCGAGCGGAGATCGTGGCGACATCAAGGAGAGCTGTGCTACGTGGCCCTGGACTCGAGAACGAGATGGCCACGGCCGCTCTCTCTCTCTGGAGAAG IleLeuThrGluArgGlyTyrSerPheValThrThrAlaGluArgGluIleValArgAspIleLysGluLysLeuCysTyrValAlaLeuAspPheGluAsnGluMetAlaThrAlaAlaSerSerSerLeuGluLys 192 AGCTACGAGCTGCCGACGGCCAGGTGATCACCATCGGCAACGAGCGCTTCGATGCCCGAGACGCTCTTCCAGCCCTCTCATCGGTATGGAGTCGGCCGGCATCCATGAGACCCTACAACAAC SerTyrGluLeuProAspGlyGlnValIleThrIleGlyAsnGluArgPheArgCysProGluThrLeuPheGlnProSerPheIleGlyMetGluSerAlaGlyIleHisGluThrThrAsnSer 281 GATCGCTGACCGTATGCAGAGGAATACCGGCTCTGGCCCCAGCACCATGAAGATCAAGATCATCGCTCCCCCGAGCGCAAGTACTCGGTGTGGATCGGGGGCTCCATCTGGCTCGCTGTCCACCTTCCAGCAGATG IleAlaAspArgMetGlnLysGluIleThrAlaLeuAlaProSerThrMetLysIleLysIleIleAlaProProGluArgLysTyrSerValTrpIleGlyGlySerIleLeuAlaSerLeuSerThrPheGlnGlnMet 309 TGGATCCACCAAGCAGGAGTACGACGAGCCGGGCCCTCCATCGTGCACCGCAATGCTTCTAG TrpIleThrLysGlnGluTyrAspGluAlaGlyProSerIleValHisArgLysCysPheEnd 375
Aldolase A	CCGGGACATCGCCAGAGGCTGCACTCCATCGGCACCGAAGACCGAAGAACACCGG GlySerIleAlaLysArgLeuGlnSerIleGlyThrGluAsnThrGluGluAsnArg 37 55 CCGGGGCCCGGCCAGGGAGTCCCTCTTCACTCTAACCCACGCTACTGA GlyAlaAlaAlaSerGluSerLeuPheIleSerAsnHisAlaTyrEnd 347 361
Glyceraldehyde 3-Phosphate Dehydrogenase	GATCCATTGATGACCTCACTACATGGTCTACATGTTCCAGTATGATTCACCCACGGCAAGTCCACGGCAAGGTCAAGGCTGAGAAGCGGAAAGCTGGTCATCAACGGGAAGGCCATCACCATTCTCCAGGAGATC AspProPheIleAspLeuHisTyrMetValTyrMetPheGlnTyrAspSerThrHisGlyLysPheHisGlyThrValLysAlaGluAsnGlyLysLeuValIleAsnGlyLysAlaIleThrIlePheGlnGluArgAsp 32 78
Myosin Heavy Chain	ACGAGGAAGTGGCAAGCAAGAGCTCTGGATGCCAGTGAACGTGTGACGCTCTCCACTCAGAATACCAAGCCTGATC SerArgLysValAlaGluGlnGluLeuLeuAspAlaSerGluArgValGlnLeuLeuHisThrGlnAsnThrSerLeuIle 1643 1669
LC2 Myosin Light Chain	GATCCAGGATTCAGGAGGCTTCAACGTCATCGATC IleGlnGluPheLysGluAlaPheThrValIleAsp 26 37 AAGGGCAACAACGAAGACGTTCTGGAGGAGCTGCTGACCACGCAAGTGTGACGCTCTCCCAAGGAGATCAAGAACATGTGGCGGCCCTTCCCCCGG LysGlyThrIleLysLysGlnPheLeuGluGluLeuLeuThrThrGlnCysAspArgPheSerGlnGluIleLysAsnMetTrpAlaAlaPheProPro 112 145
Phosphorylase	CCGGCTGCCACCTCTCGAGCAGATCTCCACGCGGGCAGCGGCTCCGG ProAlaAlaAspLeuSerGluGlnIleSerThrAlaGlyThrGluAlaSer 657 673 TCGACCAAGAGGGTACAACGCCAGGAGTACTACGACCGCATCCCGGAGCTTCGGCAGATCATCGACAGCTGAGTAGCGGCTTCTTCTCCCGAAGCAAGCCGG AspGlnArgGlyTyrAsnAlaGlnGluTyrTyrAspArgIleProGluLeuArgGlnIleIleGluGlnLeuSerSerGlyPhePheSerProLysGlnPro 721 754 CCGGTTCAAAGTCTTTCGAGATTATGAAGAGTACGTCAAGTCCAGGAGCGAGTCAAGCGCTGTACAAGAACCCCAAGAGTGGACGCGGATGGTATCCGG ArgPheLysValPheAlaAspTyrGluGluTyrValLysCysGlnGluArgValSerAlaLeuTyrLysAsnProArgGluTrpThrArgMetValIleArg 769 802
Triose Phosphate Isomerase	GATCCCAAGATTGCTGTGGCTGCACAACTGCTACAAGTGAATAACGGGCGCTCACTGGGAGATCAAGCCCTGGCATGATC AspProLysIleAlaValAlaAlaGlnAsnCysTyrLysValThrAsnGlyAlaPheThrGlyGluIleSerProGlyMetIle 56 83
α -Tropomyosin	CAGGAGCGTCTCGCAACAGCCTTGCAGAACTGGAGAGGGCGGAGAGGCGAGCAGCAGAGTCAAGAGGATGAAAGTCAATGAAAGCCGAGCCCAAGAGTGAAGAAAAATGAAATTCAGAGATC GlnGluArgLeuAlaThrAlaLeuGlnLysLeuGluGluAlaGluLysAlaAlaAspGluSerGluArgGlyMetLysValIleGluSerArgAlaGlnLysAspGluGluLysMetGluIleGlnGluIle 103 146
β -Tropomyosin	TCGAAAACCGGCCATGAAGATGAGAGAGATGGAGCTGCAAGAGATGCAAGTGAAGGAGCCCAAGCATTGCCGAGCATCAAGCCAAATACAGAGAGTGGCCAGGAAGCTGGTATCCTCGA GluAsnArgAlaMetLysAspGluGluLysMetGluLeuGlnGluMetGlnLeuLysGluAlaLysHisIleAlaGluAspSerAspArgLysTyrGluGluValAlaArgLysLeuValIleLeu 131 172
Troponin C (fast muscle)	TCGAGGAGTCTTGGTCAATGCTGCGCCAGATGAAGAGGACGCCAAGGGCAAGAGGAGGAGGATGGCCGAGTGCTTCGGCATCTCGA GluGluPheLeuValMetMetValArgGlnMetLysGluAspAlaLysGlyLysSerGluGluGluLeuAlaGluCysPheArgIlePhe 73 102 TCGAATCCCTGATGAAGGACGGGACAAAGAACACGATGGCCGATTCGACTCGA GluSerLeuMetLysAspGlyAspLysAsnAsnAspGlyArgIleAspPhe 132 148
(slow muscle)	TCGAGGAGTCAAGGACGGGCGCAAGAACACGACGGCCGATCGA GluGluLeuMetLysAspGlyAspLysAsnAsnAspGlyArgIle 134 148
Troponin I	ATGGGAGATGAAGAGAAACGCAACAAGGCCATCACGGCCCGCGGCAACCTGAAGAGCGTGATGCTGCAATC GlyAspGluGluLysArgAsnArgAlaIleThrAlaArgArgGlnHisLeuLysSerValMetLeuGlnIle 1 24 TCGAGGAGAAATCGGCATGGAAGGCCGCAAGAGATGTTCTGA GluGluLysSerGlyMetGluGlyArgLysLysMetPhe 162 174
Troponin T	GATCCCGAAGAGAGAAAGTAGTACCTTCGATGACATCCAGAAAGAGCGGCAAGAACAGGACCTCATGGAGCTGCAAGCGCTCATCGACAGCCACTCGAGGCCCGCAAGAGGAGGAGGAGCTGGTGGCCCTCAAGGAG IleProGluGlyGluLysValAspPheAspAspIleGlnLysLysArgGlnAsnLysAspLeuMetGluLeuGlnAlaLeuIleAspSerHisPheGluAlaArgLysLysGluGluGluLeuValAlaLeuLysGlu 49 AGCATCGAAGAGCCCGGCCGCAAGAGCGGAGCAAGAGGATCCGGGCCGAGAAAGGAGCGGCGGCAAGACGACTGGCGGAGGAGAGGCCCGGAGAG ArgIleGluLysArgArgAlaGluArgAlaGluGlnGlnArgIleArgAlaGluLysGluArgGluArgGlnAsnArgLeuAlaGluGluLysAlaArgArg 129 CCGGGCCAGGGTCGAGATGCTGGCCAAAGTTCAGCAAGAGGCTGGCACCAGGCCAAGGGCAAGGTCGGCGGGCGCTGGAAGTGA ArgAlaArgValGluMetLeuAlaLysPheSerLysLysAlaGlyThrThrAlaLysGlyLysValGlyGlyArgTrpLysEnd 233 259
(new isotype)	TCGAGTTCGGGAGAGCTGAAACGCCAGAAATACGACATCACCACCTTAGGAGCCGATCGA GluPheGlyGluLysLeuLysArgGlnLysTyrAspIleThrAsnLeuArgSerArgIle 217 236

Nucleotide and amino acid sequences of portions of the rabbit muscle proteins. Numbers below the sequences indicate the positions of the amino acids from the amino-terminus. (In the case of troponin C from slow skeletal muscle, the amino acids have been renumbered.)

Table 3 Sequence of cDNA clones encoding portions of troponins C and T

Troponin C

Glu
 TCGAGGAGCTCATGAAGGACGGCGACAA GAACAACGACGGCCGCATCGA
 TCGAATCCCTGATGAAGGACGGGACAA GAACAACGATGGCCGCATTGACTTCGA
 GluSerLeuMetLysAspGlyAspLysAsnAsnAspGlyArgIleAspPhe
 132 148

Troponin T

Thr Leu Ser Ile
 GACATCACCAACCTTAGGAGCCGCGATCGA
 CCGGCCAGGGTCGAG
 AspIleMetAsnValArgAlaArgValGlu
 228 237

The two nucleotide sequences and the encoded polypeptide are shown. Differences in nucleotides are denoted by overbars; when the encoded polypeptide sequence differs from that determined from the protein^{7,10}, an arrow shows the difference. The upper troponin C sequence is for slow skeletal muscle while the lower sequence is for fast muscle^{7,9}. Additional troponin T nucleotide sequences for both isotypes are shown in Table 2.

muscle proteins. Except for heavy chain myosin²⁹, there have been no previous reports of cDNA sequences for any rabbit muscle proteins and the sequences shown in Table 2 are the first to be obtained for each of the respective proteins. As mentioned above, the cDNA inserts in M13 phage are of sufficient length to use as hybridization probes for selecting full-length cDNAs from libraries where the entire cDNA is cloned directly into a vector (for example, pBR322), rather than first cut into smaller segments for the M13 sequencing strategy. We have demonstrated the use of these shorter sequences for hybridization in preliminary experiments.

The approach also tackles directly the question of isotypes, long an issue of interest, especially in the case of the troponins. Note that, if nucleotide sequence differences between isotypes are small, a cDNA clone for one could be used as a hybridization probe to obtain the others. However, the constraint of extended homology may severely limit the range of potential isotypes which can be obtained by cross-hybridization. While 8 of the 49 aligned nucleotides differ in the troponin C isotype sequences, half (7 of 15) differ in the overlapping part of the troponin T sequences (Table 3). Although the data base is small, there is clearly some indication here that cross-hybridization could be limited as an approach to isotype screening in this system.

We thank Klaus Biemann for use of the computer facilities (supported by NIH grant GM05472-24), Stephen Kuby for the sequence of rabbit adenylate kinase and sequences of creatine kinase, and Marshall Elzinga for the sequence of myosin heavy chain. This work was supported by a grant from the Muscular Dystrophy Association. S.D.P. is a postdoctoral fellow of the American Cancer Society.

Received 6 December 1982; accepted 10 March 1983.

- Dhoot, G. K. & Perry, S. V. *Nature* **278**, 714 (1979).
- Hitchcock, S. E. *Biochemistry* **12**, 2509-2515 (1973).
- Perry, S. V. & Cole, H. A. *Biochem. J.* **141**, 733-743 (1974).
- Cole, H. A. & Perry, S. V. *Biochem. J.* **149**, 525-533 (1975).
- Wilkinson, J. M. *Biochem. J.* **169**, 229-238 (1978).
- van Eerd, J. P. & Takahashi, K. *Biochem. biophys. Res. Commun.* **64**, 122-127 (1975).
- Collins, J. H., Potter, J. D., Horn, M. H., Wilshire, G. & Jackman, N. *FEBS Lett.* **36**, 268-272 (1973).
- Collins, J. H. *Biochem. biophys. Res. Commun.* **58**, 301-308 (1974).
- Wilkinson, J. M. *Eur. J. Biochem.* **103**, 179-188 (1980).
- Pearlstone, J. R., Carpenter, M. R., Johnson, P. & Smillie, L. B. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1902-1906 (1976).
- Messing, J. in *3rd Cleveland Symp. Macromolecules: Recombinant DNA* (ed. Walton, A.) 143-153 (Elsevier, Amsterdam, 1981).
- Romero-Herrera, A. E., Lehmann, H. & Castillo, O. *Biochim. biophys. Acta* **439**, 51-54 (1976).
- Frank, G. & Weeds, A. G. *Eur. J. Biochem.* **44**, 317-334 (1974).
- Weeds, A. G. *FEBS Lett.* **59**, 203-208 (1975).
- Enfield, D. L., Ericsson, L. H., Blü, H. E., Fischer, E. H. & Neurath, M. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1309-1313 (1975).
- Capony, J.-P., Pina, C. & Pechere, J.-F. *Eur. J. Biochem.* **70**, 123-135 (1976).
- Kiltz, H.-H., Keil, W., Griesbach, M., Petry, K. & Meyer, H. *Hoppe-Seyler's Z. physiol. Chem.* **358**, 123-127 (1977).

- Allen, G. *Biochem. J.* **187**, 545-563, 565-575 (1980).
- Allen, G., Bottomley, R. C. & Trinaman, B. J. *Biochem. J.* **187**, 577-589 (1980).
- Allen, G., Trinaman, B. J. & Green, N. M. *Biochem. J.* **187**, 591-616 (1980).
- Schwartz, R. J. & Rothblum, K. *Biochemistry* **19**, 2506-2514 (1980).
- Arnold, H.-H. & Siddiqui, M. A. Q. *Biochemistry* **18**, 647-654 (1979).
- Benoff, S. & Nadal-Ginard, B. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1853-1857 (1979).
- Ebashi, E. & Nonomura, Y. in *The Structure and Function of Muscle* (ed. Bovone, G. J.) 285-362 (Academic, New York, 1973).
- Hastings, K. E. M. & Emerson, C. P. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1553-1557 (1982).
- Ordahl, C. P., Kioussis, D., Tilghman, S. M., Ovitt, C. E. & Fornwald, J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4519-4523 (1980).
- Robbins, J., Freyer, G. A., Chisholm, D. & Gilliam, T. C. *J. biol. Chem.* **257**, 549-556 (1982).
- Katcoff, D. et al. *Proc. natn. Acad. Sci. U.S.A.* **77**, 960-964 (1980).
- Sinha, A. M. et al. *Proc. natn. Acad. Sci. U.S.A.* **79**, 5847-5851 (1982).
- Lemischka, I. R., Farmer, S., Racaniello, V. R. & Sharp, P. A. *J. molec. Biol.* **151**, 101 (1981).
- Aviv, H. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1408-1412 (1972).
- Crabtree, G. R. & Kant, J. A. *J. biol. Chem.* **256**, 9718-9723 (1981).
- Roberts, T. M. & Lauer, G. D. *Meth. Enzym.* **68**, 473-482 (1979).
- Maniatis, T., Jeffrey, A. & van deSande, H. *Biochemistry* **14**, 3787-3794 (1975).
- Maxam, A. & Gilbert, W. *Meth. Enzym.* **65**, 499-560 (1980).
- Messing, J., Crea, R. & Seeburg, P. H. *Nucleic Acids Res.* **9**, 309-321 (1981).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).

DNA duplication has resulted in transfer of an amino-terminal peptide between two mitochondrial proteins

T. A. Brown, R. Wayne Davies, R. B. Waring & J. A. Ray

Department of Biochemistry and Applied Molecular Biology, University of Manchester Institute of Science and Technology, Manchester M60 1QD, UK

C. Scazzocchio

Department of Biology, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK

The mitochondrial genome of the ascomycete *Aspergillus nidulans* possesses at least nine unidentified reading frames (URFs)¹⁻⁴. Two of these are particularly interesting in that the amino acid sequences of the derived translation products are very similar in the N-terminal regions. The first 36 residues of the *URFx* gene product² are repeated at the start of a second reading frame, *URFA3*, with only six mismatches. Despite this similarity in their amino-termini, the latter parts of *URFx* and *URFA3* are completely different. Closer examination of the nucleotide sequences at the 5' ends of the two genes shows that the conserved peptides are part of a 300-base pair (bp) duplication within the mitochondrial genome. This duplication also includes the tRNA gene for asparagine and a conserved 116-bp intergenic region.

The positions of the duplicates on the *A. nidulans* mitochondrial genome are shown in Fig. 1. The two copies are about 8 kilobases (kb) apart, separated from each other by the greater part of *URFA3*, the gene for apocytocrome *b* (*cobA*), a tRNA^{Cys} gene, *URF1*, *URF4*, and a tRNA^{Arg} gene.

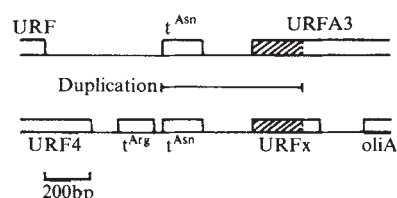


Fig. 1 Gene organization in the regions containing the upstream (top) and downstream (bottom) copies of the duplicated sequence. The homologous parts of *URFA3* and *URFx* are cross-hatched. Abbreviations: URF, unidentified reading frame; t^{Arg}, t^{Asn}, structural genes for arginine and asparagine tRNAs, respectively; *oliA*, structural gene for ATPase subunit 6. The URFs are named according to the scheme presented in ref. 4.

Fig. 2 Comparison of the nucleotide sequences of the two duplicates. Sequencing was performed by the chain termination method⁸ using the 30-bp universal primer⁹. The rationale used for sequencing the downstream copy of the duplication has been published previously². The upstream copy is located in a 1,800-bp *Bcl*I fragment of total mitochondrial DNA. This fragment was digested with *Sau*3A, and the resulting sub-fragments cloned into the *Bam*HI site of M13mp7: two of these subclones provided the duplicated sequence. The *Sau*3A site within the duplication was spanned by preparing a *Hin*II fragment for use as an internal primer onto a M13mp6 clone of the *Bcl*I fragment¹⁰. Nucleotides that are conserved between the duplicates are marked by asterisks. Comparisons of the two sequences suggest that the duplication starts either 3 or 5 nucleotides upstream of the 5' end of the tRNA genes, and ends between codons 36–37 of the unassigned genes (marked by the arrow).

```

Upstream      ...GGGAGGAGGGTATTTTAGCCTTTATAGCTCAATGGTAGAGCGGAATACTGTTAATATTTTATAGATGTTT
Downstream    ...TTATATATAAACTATTAGCCTTTATAGCTCAACGGTAGAGCGGAATACTGTTAATATTTTATAGATGTTT
               *      * *****
               |
               Start tRNAAsn
               |
AATTCATCTTAAGGGCTTATATTATAAGTTATAACTATAA---TATAGATATATATAAATTAGATCCTTAACCTAGCATAAATTAAAAA
AATTCATCTTAAGGGCTTATATTATAAGTTAAAACTATTATAGTATAGTTATATATAAATTACTACCTTAACAGCATAAATTAAAAA
*****
               |
               Start URF
               |
TAGCAAGAGAGCTATTAGAAAAAAATTAT-AAATATAAAACATGTCACAACTAATACTATTCTTTTGTAAATCAAGTAATATTT
TAGCAAGAGAGCTATTAGAAAAATTATAAGTTAATATAAAATATGCCAACACTAGTACCATTCTTTTGTAAATCAAGTAGTATTT
*****
               |
               ↓
GCATTTATAGTGTTAACTATATTAATCTATGCATTTAATAAATATATATTACCTAGATTATTATATATATACAA...
GCATTTATAGTATTAACGTATTAATCTATGCATTTAGTAAATATATATTACCTAGATTATTACGTACATATATAT...
*****

```

The nucleotide sequences of the duplicates are aligned in Fig. 2. The duplication comprises three distinct genetic units: the tRNA^{Asn} gene (nucleotide homology 99%), an intergenic spacer (homology 85%), and the open reading frame (nucleotide homology 94%, derived amino acid homology 83%; Fig. 3.) Each reading frame continues downstream of the conserved sequence, resulting in a *URFA3* translation product with a molecular weight of about 42,000, and a *URFx* polypeptide of molecular weight ~5,300. The fact that the homology between the intergenic spacers has decayed at more than twice the rate of the URFs indicates that a selective pressure exists to maintain the sequences at the N-termini of *URFA3* and *URFx*. One possible reason for this selective pressure is suggested by genetic information concerning *URFx* homologues in other organisms. Although the similarity between *URFx* and the mammalian *URFA6L* has been previously noted^{2,3}, it has only recently become known that *Saccharomyces cerevisiae* also possesses a gene having clear sequence homology to *URFx*^{5,6}. A comparison of the genome maps of *A. nidulans*, *S. cerevisiae* and *Homo sapiens* shows that in each species the unassigned gene is immediately followed by the structural gene for subunit 6 of the mitochondrial ATPase, called *oliA* in *A. nidulans* and *oli2* in yeast (Fig. 4). An association between the two genes is evident from studies with mutants of *S. cerevisiae* that map in the *URFx* homologue, *aap1*. These mutants synthesize reduced

levels of the *oli2* translation product and do not incorporate subunit 6 into the ATPase⁵, indicating that the *aap1* gene product has a role either in regulation of *oli2* expression or in assembly of subunit 6 into the ATPase. It seems probable, in view of the 50% amino acid homology between *URFx* and *aap1*, that a similar relationship exists between the URF and the subunit 6 gene in *A. nidulans*. Furthermore, as 75% of *URFx* is duplicated at the start of *URFA3*, it is likely that whatever function is performed by *URFx* with respect to *oliA* is also displayed by the amino-terminus of *URFA3* with respect to the rest of that gene. We suggest that the selective pressure that has maintained the homology between the duplicated peptides is due to the presence of a functional domain within this N-terminal region of *URFA3* and *URFx*. Although the part played by this domain cannot be deduced at present, a role in membrane attachment seems possible.

The two tRNA genes also display a very low rate of mutational decay, again indicating that both copies are genetically functional. Although iso-accepting tRNAs are not unprecedented in non-mitochondrial systems, an alternative in this genome is that primary and/or secondary structures within this segment act as transcription or RNA processing signals, as suggested for tRNA genes in mammalian mitochondrial DNAs.

The type of event that resulted in the DNA duplication cannot be inferred from the nucleotide sequence. The borders of the duplicates lack short repeated units, such as those characteristic of transposed elements, in contrast to a G+C-rich insert in the cytochrome oxidase subunit 1 gene which is bounded by 5-bp repeats (R.B.W. *et al.*, unpublished data). The downstream copy of the 300-bp duplication is equivalent in position to a segment known to be amplified in the 'ragged' growth mutant of *Aspergillus amstelodami*⁷. Amplification of this region could conceivably occur in *A. nidulans*, providing a DNA sequence that could possibly be reintegrated into another part of the genome. Alternatives to reintegration of amplified DNA are long distance intramolecular recombination or information transfer involving an RNA intermediate.

We thank Dr Phillip Nagley for communicating unpublished results concerning *aap1*. This work was supported by MRC grants to R.W.D. and C.S.

Received 22 December 1982; accepted 2 March 1983.

1. Davies, R. W. *et al.* in *Mitochondrial Genes* (eds Slonimski, P., Borst, P. & Attardi, G.) 405–410 (Cold Spring Harbor Laboratory, New York, 1982).
2. Grisi, E., Brown, T. A., Waring, R. B., Scazzocchio, C. & Davies, R. W. *Nucleic Acids Res.* **10**, 3531–3539 (1982).
3. Netzker, R., Köchel, H. G., Basak, N. & Kuntzel, H. *Nucleic Acids Res.* **10**, 4783–4794 (1982).
4. Brown, T. A., Davies, R. W., Ray, J. A., Waring, R. B. & Scazzocchio, C. *EMBO J.* (in the press).

```

URFA3:  MSQILFFFNQVIFAFIVLTILYAFNKYILPRLLYIYTHHKVKNKV...
URFx:   MPQLVPFFFNQVVFVAFIVLTVLIAFSKYILPRLRTYISRIYINKL
      * * * * *

```

Fig. 3 Comparison of the derived amino acid sequences of *URFA3* and *URFx*. The sequences are written in the International Union of Biochemistry one-letter code with the homology breakdown point shown by the arrow. Conserved residues are indicated by asterisks. Only the amino-terminal regions of the gene products are shown.



Fig. 4 Comparison between the mitochondrial genome maps of *A. nidulans* (a), *S. cerevisiae* (b)⁶ and *H. sapiens* (c)¹¹ at the start of the ATPase subunit 6 genes. The closed boxes represent *URFx* and its homologues, with the ATPase subunit 6 genes shown as open boxes. In *H. sapiens* *URFA6L* overlaps the subunit 6 gene by 46 bp.

5. Macreadie, I. G. *et al. Biochem. Int.* **5**, 129–136 (1982).
6. Novitski, C. E. *et al. in Manipulation and Expression of Genes in Eukaryotes* (eds Nagley, P., Linnane, A. W., Peacock, W. J. & Pateman, J. A.) 257–268 (Academic, Sydney, 1983).
7. Lazarus, C. M. & Kuntzel, H. *Curr. Genet.* **4**, 99–107 (1981).
8. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
9. Anderson, S., Gait, M. J., Mayol, L. & Young, I. G. *Nucleic Acids Res.* **8**, 1731–1743 (1980).
10. Waring, R. B. *et al. Cell* **27**, 4–11 (1981).
11. Anderson, S. *et al. Nature* **290**, 457–465 (1981).

Xyloadenosine analogue of (A2'p)₂A inhibits replication of herpes simplex viruses 1 and 2

Deborah A. Eppstein, Jim W. Barnett
& Y. Vivienne Marsh

Institute of Bio-Organic Chemistry, Syntex Research, Palo Alto,
California 94304, USA

G. Gosselin & J.-L. Imbach

University of Science and Technology of Languedoc, Bio-Organic
Laboratory, ERA-CNRS No. 948, 34060 Montpellier Cedex, France

Molecules of the structure ppp(A2'p)₂A containing a 2'→5' phosphodiester bond, commonly abbreviated as 2-5A, are synthesized in interferon-treated virally-infected cells¹ and have been implicated in several systems as contributing to interferon's antiviral activity^{2–7}. The 2-5A binds to and subsequently activates an endogenous endonuclease, ultimately resulting in degradation of RNA^{3–5,8–10}. We have been interested in the use of 2-5A analogues to achieve antiviral activity without the use of interferon. For this approach to be successful, analogues must be synthesized with an increased stability (native 2-5A is rapidly degraded by cellular phosphodiesterases^{11–14}) and with increased ability to enter intact cells. Removal of the highly-negative charged 5' terminal phosphates from ppp(A2'p)₂A results in formation of the 'core' species, (A2'p)₂A, which should be able to penetrate intact cells more readily. While Kimchi *et al.* have shown that 2-5A core has an antimitogenic effect in mouse spleen lymphocytes^{15,16} and 3T3 fibroblasts¹⁷, Williams and Kerr have reported lack of antiviral activity against Semliki Forest virus or encephalomyocarditis virus by exogenously-administered 2-5A core³. We have previously determined that (xyloA2'p)₂xyloA (abbreviated as xylo 2-5A core), the xyloadenosine analogue of the 5'-terminally dephosphorylated 2-5A core, is over 100 times more stable than the parent 2-5A core species¹⁸. We now report that this xylo 2-5A core inhibits replication of herpes simplex viruses 1 and 2 *in vitro*, with greater than 100 times the activity of the parent 2-5A core. The mechanism of antiviral action of the 2-5A core analogue appears to involve a pathway different from that activated by the parent 5' triphosphorylated 2-5A species.

Antiviral activities of xylo 2-5A core (synthesized as previously described¹⁹) and the parent 2-5A core (PL Biochemicals) against HSV-1 in murine 3T3 fibroblasts were determined as described in the legend to Table 1. Both compounds were determined to be >99.8% pure (A₂₆₀) by reverse-phase HPLC analysis¹⁸. Even when added to the cultures pre- as well as post-infection, 100 µM parent 2-5A core had no antiviral activity against HSV-1 when virus was determined after a single cycle of virus replication (Table 1A). By contrast, xylo 2-5A core reduced single-cycle virus yield by 72% at 1 µM ($P < 0.01$), and by >99% at 10 µM ($P < 0.001$). Pretreatment of the cells was not necessary to see this reduction in HSV-1 yield by xylo 2-5A core (Table 1A, expt 2).

When cells were infected with HSV-1 at a low multiplicity of infection (MOI) (0.025 plaque-forming units (PFU) per cell) and then treated with 1 µM xylo 2-5A core, an 87% reduction in multiple-cycle virus yield was obtained (Table 1A, expt 3, $P < 0.01$). Treatment with 10 µM xylo 2-5A core inhibited virus

yield by $4 \times \log_{10}$ ($P < 0.01$). By contrast, treatment with even 100 µM parent 2-5A core only resulted in less than $1 \times \log_{10}$ reduction in virus yield ($P < 0.01$).

We also tested the xylo 2-5A core for its ability to inhibit multiplication of HSV-2. In general, HSV-2 is more difficult to inhibit than HSV-1, and often antiherpes drugs which are active against HSV-1 are less so against HSV-2²⁰. As shown in Table 1B, yield of HSV-2 was also reduced by xylo 2-5A core.

The dose of xylo 2-5A core necessary to obtain a 50% inhibition of virus growth (ED₅₀) was determined by plaque-reduction assay. For reference it was compared with that of acycloguanosine (ACG), which is a nucleoside analogue that has potent and selective antiherpes virus activity due to its specific phosphorylation by the HSV thymidine kinase²¹. Vero cells were used in this test, demonstrating that the antiviral activity was not specific for 3T3 cells. The ED₅₀ of xylo 2-5A core against HSV-1 was 1 µM, as compared with 0.25 µM with ACG (Fig. 1A). The ED₅₀ values against HSV-2 were 1.75 and 1.70 µM for xylo 2-5A core or ACG, respectively (Fig. 1B). Thus, xylo 2-5A core shows activity *in vitro* comparable to that shown by ACG in inhibiting replication of HSV-2. Xylo 2-5A core was also effective in inhibiting herpes viral replication in a human diploid fibroblast line, MRC-5, although the ED₅₀ was higher (7–10 µM).

We have previously shown that xylo 2-5A core had an inhibitory effect on mitogen-stimulated host (3T3) cellular DNA

Table 1 Virus-yield reduction of HSV-1 and -2 by xylo 2-5A core

Treatment	Concentration (µM)	Virus titre† (PFU ml ⁻¹)	% Of control‡
A, HSV-1			
Expt 1§			
Media	—	(2.0 ± 0.4) × 10 ⁵	100
(A2'p) ₂ A	100	(2.0 ± 0.4) × 10 ⁵	100
(XyloA2'p) ₂ xyloA	1	(5.6 ± 1.3) × 10 ⁴	28**
(XyloA2'p) ₂ xyloA	10	(1.1 ± 0.5) × 10 ³	0.55***
Expt 2			
Media	—	(2.3 ± 0.8) × 10 ⁴	100
(XyloA2'p) ₂ xyloA§	10	(1.6 ± 0.2) × 10 ²	0.70**
(XyloA2'p) ₂ xyloA	10	(1.7 ± 0.3) × 10 ²	0.74**
Expt 3¶			
Media	—	(1.5 ± 0.5) × 10 ⁶	100
(A2'p) ₂ A	1	(1.3 ± 0.1) × 10 ⁶	87
	10	(1.9 ± 0.6) × 10 ⁶	127
	33	(1.1 ± 0.4) × 10 ⁶	73
	100	(4.3 ± 1.1) × 10 ⁶	29**
(XyloA2'p) ₂ xyloA	1	(1.9 ± 0.5) × 10 ⁵	13**
	10	(2.6 ± 2.7) × 10 ²	0.017***
	33	(0.5 ± 0.3) × 10 ²	0.003***
B, HSV-2*			
Media	—	(4.6 ± 0.8) × 10 ³	100
(XyloA2'p) ₂ xyloA	1	(1.3 ± 1.5) × 10 ³	28
	10	(1.0 ± 1.0) × 10 ²	2*
Xyloadenosine	30	(4.0 ± 3.5) × 10 ³	87
ACG	1	(3.6 ± 2.3) × 10 ²	8*
	10	(4.6 ± 2.8) × 10 ²	10*

† Virus yield was determined 1, 2, or 2-1/2 days (A (expts 1 and 2), A (expt 3) or B, respectively) after virus adsorption, by plaque assay in Vero cells. Each titre is the average of three independent determinations, assayed in duplicate.

‡ Results were compared using Student's *t*-test, after log-transformation of the data such that variances were comparable. *, Significant at $P < 0.05$; **, significant at $P < 0.01$; ***, significant at $P < 0.001$.

§ 3T3 cells were treated with test compounds for 5 h, washed, then infected with HSV-1, strain F, at an MOI of 5 PFU per cell. After 1 h absorption, the test compounds were re-added.

|| 3T3 cells were infected with HSV-1 (F) at 5 PFU per cell, and test compounds were added after a 1 h adsorption period.

¶ 3T3 cells were infected with 0.025 PFU of HSV-1 (F) per cell, and test compounds were added after a 1 h adsorption period of the virus.

* Vero cells were infected with 0.025 PFU of HSV-2, strain G, per cell, and test compounds were added after a 1 h adsorption period of the virus.

synthesis approximately five times greater than that obtained with parent 2-5A core, with inhibition ($\geq 90\%$) at 10 but not 1 μM concentration¹⁸. However, in order to assess more accurately the possibility that cellular cytotoxicity may contribute to the antiviral effect of xylo 2-5A core, we determined the core's effect on exponentially growing Vero cells. Rates of DNA, RNA and protein synthesis were determined by incorporation of ^3H -thymidine, ^3H -uridine or ^3H -leucine, respectively, into acid-insoluble material. 10 μM xylo 2-5A core did not appreciably inhibit cellular DNA synthesis at 24 h, although it resulted in a slight inhibition of RNA and protein synthesis (10 and 23% inhibition, respectively). However, since the ED_{50} values of xylo 2-5A core against HSV-1 and -2 were 1 and 1.75 μM , respectively, in these same cells, viral selectivity is apparent. Additional evidence that the antiviral activity of xylo 2-5A core against herpes viruses was not due to nonspecific cytotoxicity was apparent in that not all viruses tested were similarly inhibited. Thus, no antiviral activity was obtained against the RNA viruses encephalomyocarditis virus, vesicular stomatitis virus and parainfluenza-3. Xylo 2-5A core was active, however, against another DNA virus, vaccinia virus (strain WR; ED_{50} = 0.3 μM in L cells).

The activity of this xylo 2-5A core analogue against the herpes simplex viruses is remarkable. It is more than 100 times more active than the parent 2-5A core. We have previously

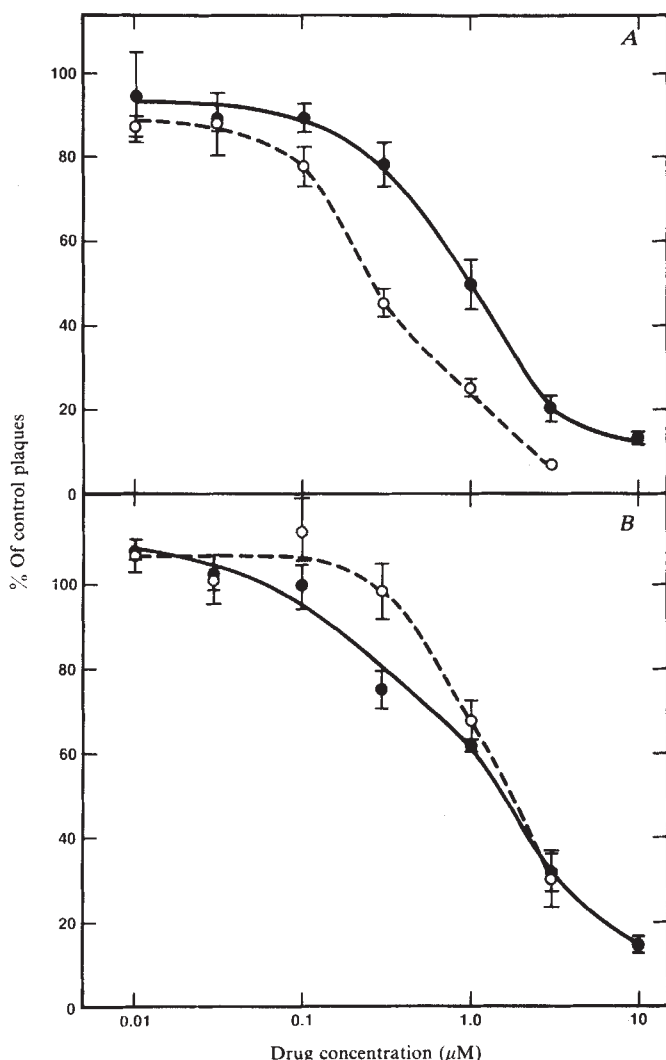


Fig. 1 Determination of ED_{50} against HSV-1 and HSV-2 of Xylo 2-5A core and acycloguanosine. Vero cell confluent monolayers (in 6-well plates) were infected with 100 PFU of A, HSV-1(F), or B, HSV-2(G). After adsorption for 1 h, agar overlay containing the indicated concentrations of drug was added to triplicate wells. After incubation for 3 days at 37°C, the number of viral plaques was counted. ●, Xylo 2-5A core; ○, ACG.

determined that this xyloadenosine 2-5A core analogue is also approximately 120 times more stable than the parent 2-5A core, with a half life of 21 h, as compared with 10 min for parent 2-5A core¹⁸. Thus, this increased stability may be contributing to the increased activity. However, we have also previously shown with other 2-5A core analogues that increased stability does not necessarily result in increased antimitogenic activity¹⁸.

The mechanism by which the xylo 2-5A core inhibits herpes simplex virus replication is intriguing. We have found that in non-infected 3T3 cells, neither the xylo 2-5A core nor the parent 2-5A core result in activation of the 2-5A dependent endonuclease (manuscript in preparation). Thus, the cores most likely do not become rephosphorylated to the 5' di- or triphosphate species within these intact cells as has been previously speculated¹⁷. The activity against several DNA, but not RNA viruses, as well as the effect on mitogen-stimulated cellular DNA synthesis, suggest that the xylo 2-5A core may be exerting its inhibitory effect through inhibition of DNA synthesis. However, we did not detect any direct inhibitory activity on virally-directed DNA synthesis in isolated nuclei from HSV-1 infected cells (data not shown).

9- β -D-xylofuranosyladenine (xyloadenosine), which is a degradation product of xylo 2-5A, showed little antiviral activity at 30 μM against HSV-2 as determined by virus yield (Table 1B). When tested in plaque-reduction assay in Vero cells, xyloadenosine was approximately 10–20-fold less active than was xylo 2-5A core in inhibiting replication of HSV-1 and HSV-2, with ED_{50} values of 20–40 μM (data not shown). De Rudder *et al.*²² have previously reported that high concentrations (33–333 μM) of xyloadenosine had some antiviral activity against HSV-1. The fact that higher concentrations of xyloadenosine than of xylo 2-5A core are required for antiherpes activity argues against the likelihood that the antiviral activity of xylo 2-5A core is directly mediated via its rapid degradation to xyloadenosine. However, these data do not rule out the possibility that a slow, sustained release of low levels of xyloadenosine monophosphate from the xylo 2-5A core may exert an antiviral activity more effectively than is obtained by a single addition of higher concentrations of xyloadenosine. Preliminary data, showing partial antagonism of the antiviral activity of both xyloadenosine and xylo 2-5A core by large excesses of adenosine, indicate that such breakdown products may indeed be contributing to the antiviral activity of xylo 2-5A core. We are now further investigating the mechanism of the antiviral activity by the parent 2-5A core and the xylo 2-5A core analogue.

We thank Carole Kurahara for cell culture support, Don Smee for performing the parainfluenza virus assays, and Julien P. H. Verheyden for helpful discussions. This is contribution no. 151 from the Institute of Bio-Organic Chemistry, Syntex Research.

Received 17 August 1982; accepted 11 February 1983.

- Williams, B. R. G., Golgher, R. R., Brown, R. E., Gilbert, C. S. & Kerr, I. M. *Nature* **282**, 582–586 (1979).
- Williams, B. R. G., Golgher, R. R. & Kerr, I. M. *FEBS Lett.* **105**, 47–52 (1979).
- Williams, B. R. G. & Kerr, I. M. *Nature* **276**, 88–90 (1978).
- Eppstein, D. A. & Samuel, C. E. *Virology* **89**, 240–251 (1978).
- Baglioni, C., Minks, M. A. & Maroney, P. A. *Nature* **273**, 684–687 (1978).
- Nilsen, T. W., Wood, D. L. & Baglioni, C. *Nature* **286**, 176–180 (1980).
- Nilsen, T. W., Maroney, P. A. & Baglioni, C. *J. Virol.* **42**, 1039–1045 (1982).
- Ratner, L. *et al. Biochem. biophys. Res. Commun.* **81**, 947–954 (1978).
- Clemens, M. J. & Williams, B. R. G. *Cell* **13**, 565–572 (1978).
- Ball, L. A. & White, C. N. *Virology* **93**, 348–356 (1979).
- Eppstein, D. A., Peterson, T. C. & Samuel, C. E. *Virology* **98**, 9–19 (1979).
- Minks, M. A., Benveniste, S., Maroney, P. A. & Baglioni, C. *Nucleic Acids Res.* **6**, 767–780 (1979).
- Williams, B. R. G., Kerr, I. M., Gilbert, C. S., White, C. N. & Ball, L. A. *Eur. J. Biochem.* **92**, 455–462 (1978).
- Schmidt, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, 4788–4792 (1979).
- Kimchi, A., Shure, H. & Revel, M. *Nature* **282**, 849–851 (1979).
- Kimchi, A., Shure, H. & Revel, M. *Eur. J. Biochem.* **114**, 5–10 (1981).
- Kimchi, A. *et al. FEBS Lett.* **134**, 212–216 (1981).
- Eppstein, D. A. *et al. J. biol. Chem.* **257**, 13390–13397 (1982).
- Gosselin, G. & Imbach, J. L. *Tetrahedron Lett.* **22**, 4699–4702 (1981).
- De Clercq, E. *et al. J. infect. Dis.* **141**, 563–574 (1980).
- Elion, G. B. *et al. Proc. natn. Acad. Sci. U.S.A.* **74**, 5716–5720 (1977).
- De Rudder, J., Andreff, F. & De Garihe, M. P. *C. r. hébd. Seanc. Acad. Sci., Paris* **264D**, 677–680 (1967).

Conjugative plasmids in bacteria of the 'pre-antibiotic' era

Victoria M. Hughes & Naomi Datta

Department of Bacteriology, Royal Postgraduate Medical School, Du Cane Road, London W12 0HS, UK

Antibiotic resistance is common in bacteria that cause disease in man and animals and is usually determined by plasmids. The prevalence of such plasmids, and the range of drugs to which they confer resistance, have increased greatly in the past 25 yr. It has become clear from work in many laboratories that plasmids have acquired resistance genes, of ultimately unknown origin, as insertions into their circular DNA. The intensive use of antibiotics since their introduction in the 1940s can explain the spread of plasmids that have acquired such genes but little is known of the incidence of plasmids in pathogenic bacteria before the widespread use of antibiotics in medicine. E. D. G. Murray collected strains of Enterobacteriaceae from 1917 to 1954; we now report that 24% of these encode information for the transfer of DNA from one bacterium to another. From at least 19% of the strains, conjugative plasmids carrying no antibiotic resistance were transferred to *Escherichia coli* K-12.

The cultures collected by Murray were contained in nutrient agar in heat-sealed glass tubes kept at room temperature. About 70% were viable when opened in 1980. They consisted of strains isolated by Murray and strains from other laboratories, almost all from human infections, and came from widely separated areas including Europe, Malta, the Middle East, northern Russia, India and North America. The accompanying notes showed that there were many duplicates—that is, the same isolate was represented more than once, sometimes as colonial variants. We tested 692 strains, of which 433 were separate isolates although some were epidemiologically related. The bacterial genera and the years of collection are shown in Table 1. Strains received by Murray from other collections had been isolated earlier, for example the type species *E. coli* isolated by T. Escherich about 1885. Our tests were to identify bacterial characters likely to be plasmid-determined: colicinogeny; haemolysis; resistance to antibiotics, other antibacterial drugs

and inorganic salts; and ability to mobilize non-conjugative resistance plasmids. We did not look for plasmids determining characters used in strain identification such as Lac⁺, H₂S production, citrate utilization, or for those associated with virulence.

Conjugative function was shown by direct transfer of resistance to rifampicin-resistant *E. coli* K-12 J53-2 or by mobilization and transfer of R300B or pHH1310a, or both, in triple crosses¹. Plasmid R300B (also called RSF1010)², molecular weight (MW) 5.7×10^6 , determines resistance to streptomycin and sulphonamides. It is easily mobilized for transfer by many conjugative plasmids, but not by plasmid F or related resistance plasmids³. We tested it for mobilization by using standard plasmids of different incompatibility groups, that is, plasmids having different transfer genes and conjugative pili⁴. Plasmid pHH1310a, MW 7.4×10^6 , determines ampicillin resistance and is compatible with R300B and with conjugative plasmids of all known groups. It was known to be mobilizable by plasmids of the IncF complex⁵. Table 2 shows comparative results of mobilization experiments with R300B and pHH1310a. One or both of the indicator plasmids were mobilized by every standard plasmid tested.

The results of our tests on the Murray strains are shown in Table 1; colicinogeny (Col) was demonstrated in 14%. Of the 36 Col plasmids transferred to *E. coli* K-12, 35 determined colicin I. Of these, 15 came from strains of *Shigella sonnei* isolated in 1937 and 10 from *Salmonella* (of unnamed species) isolated in 1941, which may mean that they originated from outbreaks of infection in those years. Only two of the cultures were haemolytic; after mating with each strain, no transfer of haemolysin was detected in 10^4 colonies of J53-2 on horse blood agar incorporating rifampicin, nor did either strain mobilize a non-conjugative plasmid. We believe therefore that haemolysis was an unstable character. There were duplicate *E. coli* cultures, both labelled as haemolytic, of which one was non-haemolytic, the other a mixture of haemolytic and non-haemolytic clones. In each, a single plasmid of MW 57×10^6 was detected after gel electrophoresis. This was the only *E. coli* strain giving haemolytic colonies but according to Murray's notes 18 of the 32 *Escherichia* isolates were originally haemolytic.

We found little antibiotic resistance, none other than was probably normal for the particular species. Only one of 30 *Klebsiella* strains was resistant to ampicillin, although most

Table 1 Characters of the Murray collection of enterobacteria

Genus	No. of strains	Dates	Colicinogeny	Haemolysis	Resistance to:						Strains having conjugative ability; pMUR established in <i>E. coli</i> K-12	
					Ampicillin	Tetracycline	HgCl ₂	K ₂ TeO ₃	CuSO ₄	NaAsO ₂	—	+
<i>Salmonella</i>	210	1917–54	15 (15)	0	0	0	0	1	39	19	6	41
<i>Shigella</i>	140	1917–52	31 (21)	0	0	0	0	0	10	15	6	29
<i>Escherichia</i>	32	1930–41	10	1	0	0	1	1	4	10	2	6
<i>Klebsiella</i>	30	1931–49	3	0	1	0	1	2 (1)	4	2	3	6
<i>Proteus</i>	16	1919–40	0	0	0	9	1	6	10	11	3	0
Unidentified	5	1920–42	0	1	1	0	0	1	1	4	0	2
Total	433		59 (36)	2	2	9	3	11 (1)	68	61	20	84

Values in parentheses are the numbers that transferred the respective character to *E. coli* K-12. pMUR indicates a conjugative plasmid from a Murray strain. Cultures are listed under the genera indicated by Murray's notes, which were confirmed by a few biochemical tests and slide agglutinations. Further systematic studies will be made at the National Collection of Type Cultures, London. The 433 cultures were all separate isolates. When the same isolate occurred two or more times in the collection, the character listed was not always found in each example. The dates are the times when cultures were added to the collection, usually the date of isolation but a few had been isolated earlier. Colicinogeny was recognised by overlaying chloroform-killed colonies with the indicators *E. coli* K-12 Row and colicin I-resistant BC3; haemolysis was identified by growing colonies on horse blood agar; and drug resistance was identified as growth across agar ditches incorporating the relevant substance, each at a concentration to inhibit the control, *E. coli* K-12. None of the strains was resistant to streptomycin, chloramphenicol, kanamycin, gentamicin, rifampicin, sulphonamide, trimethoprim or silver nitrate. When a strain appeared resistant by the ditch test, confirmation was provided by culturing the strain on whole plates incorporating the substance; concentrations of antibiotics were as used previously²⁰. Inorganic salts were present at levels that allowed growth of strains having known resistance plasmids; R478 (10^{-4} M mercuric chloride, 10^{-4} M tellurite)¹⁰, R773 (4×10^{-5} M arsenite)², pMG101 (10^{-3} M silver nitrate)²¹ and pRJ1004 (7.5×10^{-3} M copper sulphate) (T. J. Tetaz and R. K. J. Luke, personal communication). Conjugative ability was detected by transfer of resistance to *E. coli* K-12 J53-2 (*pro*, *met*, *Rif*)—either resistance of the Murray strain itself or mobilization of a non-conjugative resistance plasmid, R300B (streptomycin and sulphonamide) or pHH1310a (ampicillin). Broth cultures were mixed on the surface of nutrient agar plates. Strains resistant to any test substance were mixed with J53-2, and all strains were mixed in a three-way cross with J62 (*pro*, *his*, *trp*, *lac*) carrying R300B, pHH1310a and J53-2. After overnight incubation at 37 °C the mixed growth was taken up and plated onto minimal salts agar incorporating glucose, proline, methionine, rifampicin and the appropriate drug or salt to select resistant transconjugants. The inoculum was $\sim 10^{10}$ cells of each component evenly spread over half of the plate, then streaked over the other half. Colonies from these selection plates were purified, identified and examined for (1) the ability to transfer resistance to another K-12 strain and (2) plasmid bands, after agarose gel electrophoresis of lysates²². If transconjugants did not transfer resistance, further colonies from the selection plates were tested for resistance transfer.

Table 2 Mobilization of two non-conjugative plasmids by standard plasmids of known incompatibility groups

Mobilization efficiency	Representative plasmids of incompatibility group mobilizing:	
	R300B	pHH1310a
++	B, I, M, P, W, X	B, I, M, P, W, X, N, T, HII
+	N	C, D, F, K
±	D, HII	HI, J
ND	C, F, HI, J, K, T	

ND, none detected. F and I indicate plasmids of the F or I complex. Standard plasmids were as listed in refs 2, 4 and 10. The method was as for the Murray strains (Table 1) except that instead of a wild strain, a distinguishable K-12 strain carrying a known plasmid was used. Confluent growth on selection plates was scored as ++, semiconfluent as + and separate colonies as ±. Counts from sample matings showed that ++ represented transfer frequencies $>10^{-4}$ per donor, + $\sim 10^{-5}$ and ± $\sim 10^{-8}$.

present strains of *Klebsiella* isolated from human infections are ampicillin-resistant, having chromosomally determined β -lactamases⁶. This need not imply that *Klebsiella* as a genus has acquired ampicillin resistance since 1949; most of the *Klebsiella* species isolated by Murray were recorded by him as 'Friedländer's bacillus' and they may prove to belong to different species (or biotypes) from the *Klebsiella* species prevalent today. The *Klebsiella* species differ in their β -lactamases and naturally resistant biotypes may have been selected in the recent past. The only tetracycline-resistant cultures belonged to the genus *Proteus*, most strains of which have chromosomally determined tetracycline resistance. We found no other antibiotic resistance and no resistance to sulphonamides, although the collection extended well into the era of sulphonamides and into the early years of antibiotic therapy.

A high proportion of the present day resistance plasmids determine mercuric chloride resistance^{7,8} for which there is no satisfactory explanation. Only three of the Murray strains showed this resistance and none transferred it to *E. coli* K-12. The only transferable resistance plasmid we identified determined tellurite resistance, which is relatively uncommon in antibiotic resistance plasmids^{9,10}. Resistance to copper sulphate and sodium arsenite was a frequent finding but no transfer of either resistance was demonstrated. Although these resistances may be plasmid-borne¹¹⁻¹³ there is no evidence that they are commonly so. They occur sufficiently frequently in *E. coli* strains to have been used in epidemiological typing¹⁴; we have no information on their general occurrence in bacteria other than *E. coli*.

Gene transfer was determined by 104 strains; one strain showed transfer of tellurite resistance and 103 strains demonstrated the ability to mobilize a non-conjugative plasmid. Their dates of collection were from 1917 to 1941. Conjugative plasmids from 84 strains were transferred to and established in *E. coli* K-12; 20 other strains brought about transfer of R300B or pHH1310a but genetic determinants for mobilization were not detected (10 cases) or only transiently present (10 cases) in K-12. Of the 84 plasmids in K-12, 36 determined colicinogeny, one determined tellurite resistance and 47 were cryptic. All the ColI plasmids had molecular weights between 5.5×10^7 and 6×10^7 , measured by comparison with reference plasmids of known molecular weights after agarose gel electrophoresis of K-12 lysates. The tellurite resistance plasmid had a MW of 1.05×10^8 and the other conjugative plasmids, except for two that were undetected in agarose gels, were in the range $2.5-7.0 \times 10^7$. The indicator plasmids, R300B and pHH1310a, were visible in the agarose gels. In addition, lysates of K-12 strains that had acquired conjugative plasmids from cultures of *S. sonnei*, but not from other species, revealed cryptic plasmids of $<5 \times 10^6$. Small cryptic plasmids in antibiotic-resistant strains of *S. sonnei* were reported by Jamieson *et al.*¹⁵.

Soon after genetic recombination in bacteria was discovered, among 2,000 wild strains of *E. coli* screened for the ability to transfer chromosomal genes to *E. coli* K-12 F⁻, 2% gave positive results¹⁶. This represents, as far as we know, the only relevant survey of cultures of the same period as the Murray collection and used, of course, a much less sensitive test for genetic transfer than a plasmid mobilization test¹. The latter test has been used more recently to identify conjugative plasmids in drug-sensitive bacteria and their prevalence was found to be comparable with that in the Murray collection, that is, 17% of 300 faecal *E. coli* strains from people outside hospital¹⁷ or 20 of 60 *E. coli* strains from healthy pigs, cattle and human beings¹⁸.

The proportion (24%) of Murray strains in which genetic transfer function was demonstrated is probably an underestimate because of plasmid loss or the use of an inappropriate screening test. We do not claim that the cultures are perfectly representative of the relevant bacterial genera over the period of the collection, but they are many and varied enough for us to conclude that conjugative plasmids were as common in enterobacteria before the medical use of antibiotics as they are in drug-sensitive strains today. This was not wholly expected as it was believed that the use of antibiotics has encouraged the spread of bacterial clones carrying such plasmids¹⁹. Plasmids that determine antibiotic resistance have a symbiotic relationship with their bacterial hosts. It is unknown whether the conjugative plasmids that we have identified encode characters useful to their hosts or whether they are entirely parasitic.

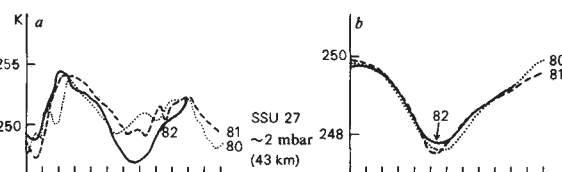
We thank Professor R. G. E. Murray for allowing us access to his father's strains, now kept at the National Collection of Type Cultures, London.

Received 6 December 1982; accepted 10 February 1983.

- Anderson, E. S. *Nature* **208**, 1016-1017 (1965).
- Jacob, A. E. *et al.* in *DNA Insertion Elements, Plasmids and Episomes* (eds Bukhari, A. I., Shapiro, J. A. & Adhya, S. L.) 607-638 (Cold Spring Harbor Laboratory, New York, 1977).
- Willets, N. & Crowther, C. *Genet. Res.* **37**, 311-316 (1981).
- Bradley, D. E. *Plasmid* **4**, 155-169 (1980).
- Datta, N., Hughes, V. & Nugent, M. in *Plasmids of Medical, Environmental and Commercial Importance* (eds Timmis, K. N. & Pühler, A.) 175-182 (Elsevier, Amsterdam, 1979).
- Matthew, M. & Harris, A. M. *J. gen. Microbiol.* **94**, 55-67 (1976).
- Schottel, J., Mandal, A., Clark, D., Silver, S. & Hedges, R. W. *Nature* **251**, 335-337 (1974).
- Nakahara, H., Ishikawa, T., Sarai, Y., Kondo, I. & Mitsuhashi, S. *Nature* **266**, 165-167 (1977).
- Summers, A. O. & Jacoby, G. A. *J. Bact.* **129**, 276-281 (1977).
- Bradley, D. E., Hughes, V. M., Richards, H. & Datta, N. *Plasmid* **7**, 230-238 (1982).
- Ishihara, M., Kamio, Y. & Terawaki, Y. *Biochem. biophys. Res. Commun.* **82**, 74-80 (1978).
- Hedges, R. W. & Baumberg, S. *J. Bact.* **115**, 459-460 (1973).
- Summers, A. O. & Silver, S. A. *Rev. Microbiol.* **32**, 637-672 (1978).
- Elek, S. D. & Higney, L. *J. med. Microbiol.* **3**, 103-110 (1970).
- Jamieson, A. F., Bremner, D. A., Bergquist, P. L. & Lane, H. E. D. *J. gen. Microbiol.* **113**, 73-81 (1979).
- Lederberg, J., Cavalli, L. L. & Lederberg, E. M. *Genetics* **37**, 720-730 (1952).
- Lewis, M. J. *Lancet* **i**, 1389-1393 (1968).
- Smith, H. W. & Linggood, M. A. *J. gen. Microbiol.* **62**, 287-299 (1970).
- Anderson, E. S. *Br. med. J.* **2**, 1289-1291 (1965).
- Datta, N., Hughes, V. M., Nugent, M. E. & Richards, H. *Plasmid* **2**, 182-196 (1976).
- McHugh, G. L., Moellering, R. C., Hopkins, C. C. & Swartz, M. N. *Lancet* **i**, 235-240 (1975).
- Broome-Smith, J. *Plasmid* **4**, 51-63 (1980).

Corrigendum

In the letter 'Stratospheric warming following the El Chichón volcanic eruption' by D. E. Parker and J. L. Brownscombe, *Nature* **301**, 406-408 (1983), the temperature scales for SSU 27 in Fig. 3a and b were wrongly labelled. The correct labelling is shown below.



Recombination and directional selection

FLEXON AND RODELL¹ observed a correlated response for increased recombination in populations of *Drosophila melanogaster* selected for DDT resistance. Although they suggested that this correlated response is due to 'hitch-hiking', there are at least three possible explanations, one involving pleiotropy and two involving hitch-hiking. The initial conditions for all of them to occur are somewhat restrictive so that it is not possible to suggest a higher *a priori* likelihood for any of them.

First, the alleles that confer resistance to DDT may have some pleiotropic effect that affects recombination. In other words, the mechanism of the resistance in this population may be such that it also influences the normal process of recombination.

Second, alleles increasing recombination may be carried along by (hitch-hike with) alleles that have resistance to DDT. In this case, the increase in recombination in itself does not affect fitness, that is, the level of recombination does not affect population fitness. Of course in another population or at another time, the alleles increasing resistance may be initially associated with alleles reducing recombination and a correlated response for reduced recombination would be observed. This hypothesis assumes that there is an initial positive association (gametic disequilibrium) between alleles increasing DDT resistance and alleles increasing recombination on both chromosomes II and III.

Third, the hypothesis advocated by Flexon and Rodell is that "recombinant genes generating genotypes favoured by selection can hitch-hike to increased allele frequencies". In this case, recombination results in the production of new gametes with an increased fitness that are associated with (in gametic disequilibrium with) an allele(s) for increased recombination. This last hypothesis is obtained from a theoretical model developed to show the potential evolutionary advantage of increased recombination. It assumes repulsion gametic disequilibrium among favourable alleles at fitness loci. Both hitch-hiking hypotheses assume tight linkage among the various genes involved.

From the data given, it appears that none of these hypotheses may be falsified and one cannot state that the results 'confirm' the last hypothesis but only that the data are consistent with all of these hypotheses.

This work was supported by NIH grant ROI HD 12731 to Glenys Thomson.

Note added in proof: It should be noted that the marker loci on chromosomes II and III, the chromosomes that showed

increased recombination, span the centromeric heterochromatin and the markers on chromosome I do not. Knowledge of the distribution of crossovers, between- or intra-arm on chromosomes II and III might tell the significance of this difference.

PHILIP W. HEDRICK

Department of Genetics,
University of California,
Berkeley,
California 94720, USA

1. Flexon, P. B. & Rodell, C. F. *Nature* **298**, 672-674 (1982).

RODELL REPLIES—Hedrick's expansion of our explanation for increased recombination frequency is both pertinent and welcomed. The possibility that high recombination rate alleles may be tightly linked to high DDT resistance alleles cannot be ruled out by our experiment. To make the distinction between coupling and repulsion gametic disequilibrium requires a system whereby both increased and decreased character expression can be selected. I have a selection experiment in progress involving sternopleural chaetae number in *Drosophila melanogaster* that will distinguish between these possibilities.

Hedrick's initial point, that high DDT resistance alleles may have a positive pleiotropic effect on recombination frequency, requires further comment. If the DDT resistance 'mechanism' is a cellular or organismal physiological phenomenon, it would seem that all chromosomes might be affected. We monitored chromosome regions encompassing 45.2, 88.0 and 16.8 map units for chromosomes I, II, and III, respectively. Because we observed significant recombination rate increases for chromosomes II and III but not for chromosome I, this explanation appears less likely than hitch-hiking.

CHARLES F. RODELL

Department of Biology,
College of St Benedict,
St John's University,
Collegeville, Minnesota 56321, USA

Leaching of nuclear waste forms

BARKATT *et al.*¹ have described the effects of radiolytic acidification of the aqueous medium on the leaching performances of various glass and ceramic waste forms, relative to their leaching performance in non-irradiated conditions in water of pH ~ 7. They conclude that the leaching performance of the ceramic Synroc-D could deteriorate seriously in active conditions.

In response, I would first like to emphasize that the above experiments have no bearing on the performance of Synroc-C, the titanate ceramic designed to immobilize the high-level waste (HLW) arising from the reprocessing of spent fuel from civilian power reactors. Several independent laboratories have confirmed that Synroc-C is far more resistant to leaching than borosilicate glass over a wide range of temperatures and pH conditions.

Synroc-D is a specialized waste form with no application to the HLW which will be derived from reprocessing the inventories of spent fuel from commercial power reactors, which are widespread in Europe, the United States and Japan. Synroc-D is a specialized waste form designed solely to immobilize the wastes from the US Defense program that are held in tank storage at Savannah River and Hanford. Unlike commercial HLW, the defense wastes contain large quantities of processing contaminants, including iron, aluminium, sodium and silica. Thus, caesium is immobilized in the aluminosilicate phase nepheline, while the remaining radionuclides are incorporated in titanate phases (as in Synroc-C). Accordingly, the ability of Synroc-D to immobilize caesium is similar to borosilicate glass, whereas multivalent elements such as uranium and rare earths are about 1,000 times less easily leached from Synroc-D than from glass.

It is unfortunate that Barkatt *et al.* did not attempt to characterize their Synroc-D sample, because some of this material now in circulation has not been fabricated according to my specifications. Thus, these samples contain in place of some or all of the nepheline, a low-silica, high-alumina glass. Some strontium and calcium as well as caesium can be found in this glass. The fact that in acid conditions, leach rates increased for Si, Al, Sr and Ca as a group¹, provides strong evidence for the presence of such a glass.

While Barkatt *et al.* have demonstrated that radiolysis can give rise to low pH conditions in the laboratory, it is important to realize that these pH levels are unlikely to prevail when a waste form is finally interred underground, because of buffering by surrounding repository rocks and canister overpacks. Aqueous solutions of pH ~ 6 would be far more realistic.

A. E. RINGWOOD

Research School of Earth Sciences,
Australian National University,
PO Box 4, Canberra 2600,
Australia

1. Barkatt, A., Barkatt, A. & Sousanpour, W. *Nature* **300**, 339-341 (1982).

BARKATT ET AL. REPLY—Ringwood points out that the large radiation-induced enhancement in leach rates reported¹ for Synroc-D has no bearing on the performance of titanate ceramics. This is in complete agreement with the major conclusion of our paper—that the radiolytic pH decrease effect leads to a considerable increase in leach rates in the case of waste-form materials which have a high alumina content. We would welcome the opportunity to include high-Ti, low-Al ceramics in future studies. In any case, Ringwood's comments confirm that an aluminosilicate phase has to be introduced in order to immobilize the Cs in defense wastes.

The Synroc-D samples used in our previous studies were prepared at Lawrence Livermore Laboratory. While there may be substantial differences in properties between these and other specimens of Synroc-D, they will indicate that the material is highly sensitive to small variations in fabrication and composition.

Measurements carried out in our laboratories² and elsewhere³ confirm Ringwood's statement that at neutral or mildly alkaline pH the leach rate of uranium from Synroc-D is substantially lower than those observed in glasses and also in Rockwell polyphase ceramics. However, in the case of Sr the Rockwell ceramics are less leachable by two orders of magnitude and in the cases of Cs and of the major matrix elements (Al and Si) high-silica glass is much less leachable. Data for some of the most hazardous species (Np, Tc) are not yet available.

The conditions which will prevail during the long service time in a geological repository are highly uncertain and liable to undergo considerable variations. In particular, water coming into contact with the waste form is likely to have a long contact time and to undergo considerable changes due to chemical (selective leaching of soluble species) and radiolytic effects. For example, the original concentration of buffering species in typical ground water (0.90 mmol l⁻¹ total CO₂ + 1.3 mmol l⁻¹ SiO₂ in basaltic Grande Ronde water) can be quickly overwhelmed by radiolytically produced acid (0.06 mmol l⁻¹ day⁻¹ total H⁺ at a moderate γ dose rate of 8.6×10^4 rad h⁻¹). Indeed, at the end of 7 days of irradiation, the pH of such water is observed to drop from 9.5 to <7 when the water is initially air-saturated. At the end of 60 days the pH is 4.

Such considerations necessitate testing in a wide variety of conditions and for a large number of components. For example, Soper *et al.*⁴ recently used a combination of tests at pH values of 3–4, 7 and 10–11 for the optimization of defense waste glass compositions. As Ringwood indicates, it is indeed possible to obtain excellent results with respect to the leachability of a single component at a single pH value (for example, U at pH

6) through careful tailoring of the waste-form composition to be used with a particular waste stream. However, as emphasized above, a primary consideration in the comparative evaluation of waste forms should be the recognition that the chemical properties of the hazardous elements in waste streams are highly diverse and that the range of possible future conditions in a repository is very broad. Accordingly, evaluation of waste-form materials should emphasize demonstration of a relatively low sensitivity of the leach rates of their various components to changes in the chemical environment rather than performance with respect to a single component in one particular set of conditions.

AARON BARKATT
ALISA BARKATT
WILLIAM SOUSANPOUR

*Vitreous State Laboratory,
The Catholic University of America,
Washington DC 20064, USA*

1. Barkatt, A., Barkatt, A. & Sousanpour, W. *Nature* **300**, 339–341 (1982).
2. Barkatt, A., Barkatt, A. & Sousanpour, W. *Nucl. Chem. Waste Mgmt* (in the press).
3. Bernadzkiowski, T. *et al.* *The Evaluation and Selection of Candidate High-level Waste Forms* (US Dept Energy Rep. DOE/TIC-11611, 1982).
4. Soper, P. D., Roberts, G. J., Lightner, L. F., Walker, D. D. & Plodinec, M. J. *Bull. Am. ceram. Soc.* (in the press).

Effect of interferon on vesicular stomatitis virus proteins

In studying the effect of interferon (IFN) on protein glycosylation¹, the cells used by Olden *et al.* were markedly less sensitive to IFN than was the line we have used. With 30 units of IFN they report a 10–20-fold inhibition of production of vesicular stomatitis virus (VSV) particles of unchanged specific infectivity. We agree that this concentration of IFN reduces the number of particles by around 10–20 fold but we regularly see a 300–600-fold inhibition of infectivity at this concentration. This almost certainly accounts for the differences seen between Olden's group and ours in the amount of VSV-associated G protein in virus produced from IFN-treated cells. With only 1/10th to 1/50th of the virus inhibition, it is most unlikely that Olden *et al.* could have seen differences in the quantity of viral G (or, also reported M) protein comparable with those originally reported².

We have reported our finding of a reduction in G and M proteins in VSV produced by IFN-treated mouse L_B and the results have been confirmed and extended by Jay³. Finally, this work on VSV produced by IFN-treated cells has also been confirmed in this and other laboratories^{4,5} with another membrane-associated virus from an entirely different group, murine leukaemia virus (MLV). These studies report that MLV deficient

in glycoprotein (gp 67/71) was produced by IFN-treated cells. Another study⁶ reported inhibition of the rate of virus protein glycosylation in IFN-treated cells that were infected with MLV.

R. M. FRIEDMAN
R. K. MAHESHWARI

*Department of Pathology,
Uniformed Services University
of the Health Sciences,
Bethesda, Maryland 20814, USA*

1. Olden, K., Bernard, B. A., Turner, W. & White, S. L. *Nature* **300**, 290–292 (1982).
2. Maheshwari, R. K., Banerjee, D. K., Waechter, C. J., Olden, K. & Friedman, R. M. *Nature* **287**, 454–456 (1980).
3. Jay, F. T., Darwood, M. D. & Friedman, R. M. *J. gen. Virol* (in the press).
4. Friedman, R. M., Maheshwari, R. K., Jay, F. T. & Czarniecki, L. *Ann. N.Y. Acad. Sci.* **350**, 533–544 (1980).
5. Bilello, J. A., Wivel, N. A. & Pitha, P. M. *J. Virol.* **43**, 213–222 (1982).
6. Aboud, M., Kimchi, R., Bakhanashvili, M. & Salzberg, S. *J. Virol.* **40**, 830–839 (1981).

OLDEN ET AL. REPLY—Friedman and Maheshwari do not dispute the major thesis of our manuscript, that IFN is not an inhibitor of glycosylation as originally claimed¹. The following quotation taken from the manuscript of Bilello *et al.*² supports our finding: "In the SC-1/MCF system, as well as in other cells we have studied, no inhibition of N-linked glycosylation in the presence of IFN has been observed. The envelope glycoprotein both from control and IFN-treated cells contains mannose, fucose and glucosamine".

KENNETH OLDEN
BRUNO A. BERNARD
WILLIE TURNER
SANDRA L. WHITE

*Howard University Cancer Center and
Department of Microbiology,
Washington, DC 20060, USA*

1. Maheshwari, R. K., Baneyce, D. K., Waechter, C. J., Olden, K. & Friedman, R. M. *Nature* **287**, 455–455 (1980).
2. Bilello, J. A., Wivel, N. A. & Pitha, P. M. *J. Virol.* **43**, 213–222 (1982).

Matters Arising

Matters Arising is meant as a vehicle for comment and discussion about papers that appear in *Nature*. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted. Neither contribution nor reply (if one is necessary) should be longer than 500 words and the briefest of replies, to the effect that a point is taken, should be considered.

Two types of fundamentalist

George M. Marsden

The Creation Controversy: Science or Scripture in the Schools.

By Dorothy Nelkin.

W. W. Norton: 1983. Pp. 242. \$16.95, £12.50.

WHY DO SO many Americans believe that creation and biological evolution are logical opposites? Why do they believe that the teaching of biological evolution is inherently a threat to Christianity and even to civilization itself? Even from traditional Christian perspectives these conclusions seem eminently escapable. Ever since the advent of Darwinism, many distinguished theologians, including some American evangelicals, have pointed out that God's providence could care for evolutionary development just as well as for any other natural process. Hence one's biological theories need have nothing to do with whether one believes that ultimately there is a creator. Moreover, even a high view of the authority of Scripture does not seem to entail that one should have to believe that the "days" of Genesis One were twenty-four-hour periods or that the exact order of creation has been revealed. So the Christian faith need not be thought to stand or fall with the latest fossil discovery. Even though it is true, as cultural historians would agree, that naturalistic evolutionary philosophies tend to undermine traditional values, it does not seem to follow in any obvious way that such moral erosion is best stopped by banning the teaching of a *biological* theory. Nonetheless, millions of Americans today think that inhibiting or neutralizing the teaching of biological evolution is a key to saving their civilization.

Dorothy Nelkin, in this updating of her *Science Textbook Controversies* (MIT Press, 1977), attempts to explain the phenomenon by sociological analysis. Her explanation is that the recently influential "creation-science" movement has flourished due to popular anxieties about science and technology. Nelkin is an astute observer and points out that this answer is not a simple one. For instance, her detailed description of the recent creation controversy includes notice that the creation-science movement includes a striking number of engineers. Adherents of creation-science, moreover, seem sincerely to profess the highest regard for science and scientific evidence. Nelkin points out, for instance, that the arguments by creation-scientists concerning science and its social and educational implications closely parallel views expressed by their most outspoken scientific opponents. Mere resentment against the dominance of scientific experts in a technological society

will not explain the phenomenon. Recognizing this, Nelkin resorts to a further explanation that is virtually a tautology. Champions of the creation-science movement are concerned about the threat of science to traditional "values" and about the "social implications of modern science and technology".

Nelkin's account, although highly valuable as a description of various state and local conflicts over science textbooks and teachings during the past decade, tells us little analytically because of two weaknesses common to much of sociology. First, is an underdeveloped sense of history. Nelkin's introductory historical overview is thin, unreliable and based on a naive acceptance of the myth of an inevitable warfare between science and religion. Certainly the volume should not be proclaimed, as it is on the dust jacket, "a history of the struggle between creationists and scientists from the nineteenth century to the Arkansas trial". Secondly, the author displays not the slightest interest in or understanding of differences among religious beliefs. They are for her all simply "values". Such a sociological approach accordingly misses the subtleties of the religious issues that must be considered to explain creation-science. Without these the analysis is akin to that of a highly intelligent art critic who is, nonetheless, colourblind.

One of the distinctions that eludes Professor Nelkin is that between proponents of the creation-science movement and other evangelical Christians who oppose only the promotion of exclusively naturalistic philosophies in the name of objectivity. In her detailed description of the 1970s controversies in California, she notices that these two groups often opposed each other; but then in the same account she calls both of them simply "creationists", thus confusing the two. In fact to understand why the anti-evolution movement is flourishing, one must include analysis of the peculiar beliefs of the creation-science movement, which has been behind almost every recent effort to "balance" the teaching of evolution with teaching of creation. Proponents of this movement, most of them associated with the Institute for Creation Research in San Diego, simply *mean* by "creation" a young Earth and explanation of geological evidence by a worldwide flood. They have appropriated the word "creationist" for this narrow view. Professor Nelkin, while

providing a valuable account of the ICR (supplemented in an appendix by the admirable analysis in Judge William Overton's ruling on the 1981 Arkansas law) perpetuates the confusion by sometimes using "creationism" in the ICR's narrow sense and sometimes in the sense of referring to any believer in divine creation.

Once the distinctions are made, one might ask why the ICR insists on such a narrow position. The answer lies not in a disdain for science but in a view of the Bible as conforming to the modern scientific standards of exact statement. Such literalistic views are sometimes related also to supposedly scientific prophetic interpretations that depend on exact numbers. The next question might be why such an eccentric view has made so great an impact on contemporary America. The success of the ICR seems related to technical proficiency in promotion so as to capitalize on the popular American enmity against evolution. "Evolution" in America has become a popular code-word for atheistic naturalism. To unearth the origin of this powerful symbolism would require looking especially at Southern reactions to modern culture after the Civil War and at fundamentalism after the First World War.

Such an analysis might explore the reasons why in America the metaphor of warfare has been overwhelmingly compelling in so many of the discussions concerning evolution and religion. The fondness of religious fundamentalists for warfare imagery is well known. As Nelkin points out, however, scientists have their own equivalent of fundamentalists. True science, they assume, is at war with all traditional religion. According to this view, a full account of the natural processes operative in the physical world would somehow exclude theism. Nelkin tells of how this view was promoted by social scientists in America in all ill-fated grade-school curriculum entitled *Man a Course of Study* (MACOS). MACOS frankly used evolutionary theory to suggest that all values are relative. Such efforts by scientists or social scientists to enlist the prestige of science in a philosophical cause only invite equally ill-conceived attacks on aspects of modern science.

Such episodes also suggest that there are real issues that the creation-scientists and their many followers are trying to address. It is not wholly illusory to believe that what they imprecisely call "secular humanism" sometimes operates as a virtual religion, or at least as an arbiter of cultural values. Moreover, some influential groups in America wish to eliminate traditional religion and its influences entirely from public life. Yet in a monopolistic state school system some values will be taught. In a pluralistic society, how does one determine what those values should be? What if these values offend the beliefs of

parents whose children are being educated? What rights do individual teachers, school districts or states have to impose particular values, whether secular or traditionally religious, on all children? None of these questions seems especially related to natural science, although it remains interesting why they have become so. In any case, creation-scientists and their ilk, despite their extremism and many scientific foibles, are drawing attention to some notable yet unresolved questions related to attaining genuine pluralism in American life and education. □

George M. Marsden is Professor of History at Calvin College, Grand Rapids, Michigan, and author of Fundamentalism and American Culture (Oxford University Press, 1980).

Gnawing doubts

Bernard Wood

Teeth: Form, Function, and Evolution.

Edited by Björn Kurtén.

Columbia University Press: 1982. Pp.383. \$50 (US), \$65 (elsewhere).

IN THE summer of 1979, fifty scientists with diverse interests in teeth gathered in Finland for the Fifth International Symposium on Dental Morphology. This volume contains the texts of the papers read at the meeting. The contributions cover a wide variety of topics and are themselves variable in quality.

The 25 papers have been classified under one of four topic headings — Ontogeny; Structure, Shape and Function; Populations; and finally Evolution — but the categorization is sometimes confusing. The editor is to be commended for preparing a volume index, but it would have been useful if he had insisted on a more uniform format for the papers. Some have an abstract and a conclusions section; regretably others have neither.

The first section on ontogeny (containing seven papers) is perhaps the strongest. Three (Kollar, Lumsden and Savara *et al.*) are essentially presentations of research methods, but they nonetheless represent important advances in this field of research. Two others, by Butler and Moss, are more reflective and Butler's contribution is a particularly stimulating essay on the development of tooth patterns.

Plants under threat

A revised version of the *British Red Data Book: Vascular Plants* has been issued by the Royal Society for Nature Conservation, supplanting the previous edition published in 1977. As well as details of each endangered species, the book includes brief accounts of how the information is compiled and of pertinent legal and scientific issues. Copies are available from RSNC, The Green, Nettleham, Lincoln, price £7 post paid.

In the next section two papers stand out: one for its strengths and the other for its weaknesses. Schwartz presents a critical examination of the problems of heterodonty and homology. He asks whether teeth are homologous, or whether it is the sites (e.g. the premaxilla) which are homologous with different teeth coming to rest in different places in different animals. He comes down in favour of the latter interpretation, and presents a developmental model to explain the differentiation of not only the basic tooth types, but also the variation within each type. Gantt's contribution, by contrast, is poorly edited and made more confusing by the lack of any labelling on an important figure.

The problem of dealing with the effects of attrition is one which particularly bedevils the work of hominid palaeontologists and the results presented in van Reenen's paper will be useful when evaluating the significance of differences in tooth size between presumed hominid taxa. However, the paper would have been more useful if it had included a table giving the percentage reduction in tooth size due to attrition for individual teeth.

Two papers in the section dealing with populations illustrate the extra difficulties scientists make for themselves, and others, when they are careless in their definition of morphological variants. Both Axelsson and Kirveskari, and Scott and Dahlberg, include an assessment of the incidence of the C6 in their studies, but by choosing to define it differently they render their results incompatible.

The final section, on evolution, contains only four papers, and one of these is really about ontogeny. Rensberger's contribution deals with Miocene aplodontid rodents, but is commendably functionally orientated and one wonders why it is not in the second section along with a similar paper by Fortelius dealing with rhinoceroses. The last paper, by Patricia Smith, has the seductive but, as it turns out, misleading title of "Dental Reduction: Selection or Drift". In the abstract she implies that in four continents there are examples of neighbouring populations in which tooth size can be related to diet, yet the only detailed evidence she cites or quotes for this comes from Australia. The promised discussion of the rival merits of selection and drift never materializes.

Interest in teeth is undergoing a modest revival and the increasing dialogue between developmental biologists and palaeontologists is particularly promising, but this volume conveys little of this excitement. There is evidence from at least one of the papers that the meeting provided the opportunity for fruitful discussion, but even so a successful meeting does not always result in the publication of a useful set of proceedings. Such, sadly, is the case with this book. □

Bernard Wood is Professor of Anatomy at The Middlesex Hospital Medical School, London.

Descent and dissent

Barry Cox

Archetypes and Ancestors: Palaeontology in Victorian London 1850-1875.

By Adrian Desmond.

Blond and Briggs: 1982. Pp.287. £15.95.

MOST scientists are uninterested in the history of their profession, perhaps tending to see earlier science as at best merely a collection of inadequate half-truths. Yet ideas, like animals, are adapted to their environment, now as then. The views of society and of current science constrain our choice of research topic and of the currently-respectable range of theories that we may prefer. The development of these theories is also often partly moulded, even perverted, by the theories of the other scientists with whom we interact.

But, if we try to learn History's lessons, we have to beware of simply accepting the victor's tales of the conflicts, which merely perpetuate their partisan view, turning the losers into mere reactionaries and the victors into selfless bringers of new light. Adrian Desmond has tried to set the record straight in his account of the battle that was fought, largely in the laboratories and lecture halls of London, over Darwin's new theory of evolution by natural selection. He has clearly shown the social, philosophical and personal differences which underlay the prolonged debate, in which Owen and Huxley were the leading protagonists.

The conservatism and obduracy of Owen and the Establishment have long been clearly highlighted. Desmond's contribution is to redress the balance by showing the ambitiousness, intolerance and inconsistency of Huxley and his supporters. He shows that they identified themselves with the increasingly powerful bourgeoisie, who saw the idea of evolution as the sharp edge of an instrument of social change. Their aim was the replacement of the Old Guard at the levers of power — not merely in general terms in Society as a whole, but also in personal advancement within science by capturing the prestigious and influential positions within the universities, museums and scientific societies.

Perhaps the most interesting result is the extent to which Huxley reacted to the opposition's theories, rather than single-mindedly merely weaving all the available evidence into the garment designed by Darwin. For example, Owen invented the Archetypes of the book's title as abstract blueprints, from which each of the successive fossil forms of each major group had been directly produced by gradually increasing degrees of divine modification. One might have expected Huxley just to re-interpret the data and to use that same

progressive series as evidence for Darwinian evolution. Instead he tried to deny its existence, suggesting that the fossil record showed, not progressive change, but "persistence" of many almost unchanged types, and that the real evolutionary changes had taken place in the "pre-geological" Cambrian. (Though this now sounds ridiculous, perhaps we should wonder whether there may not have been, to Huxley, little difference in degree between this lurch into the past and the then-recent surprising discovery of a Jurassic mammal.)

In another series of reactions, Desmond suggests that Owen's recognition of the Dinosauria as a group of giant, erectly-postured reptiles was an answer to the godless Lamarck's idea of the innate progression of life by showing that reptiles, far from progressing to greater glories since the Mesozoic, had in fact degenerated into mere flimsy, sprawling lizards. To counter Owen, Huxley then pointed to the bipedal dinosaurs and proclaimed that they did, after all, have more advanced descendants — the birds. Reluctant to accept any Huxleyan suggestion, Owen in his turn was driven to suggest that birds had instead descended from the bat-winged pterosaurs. The battle might have been less bitter had Owen been less proud and forbidding, or Huxley less impatient and vindictive. Nevertheless one could scarcely

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

have a better illustration of the blinkering effects of changing one's aim and trying to confound the opposition instead of trying to unravel the intricacies of Nature.

Adrian Desmond's book is written in clear and lively style, and is a useful sequel to Martin Rudwick's *The Meaning of Fossils* (Macdonald, 1972), for two reasons. First, it follows Rudwick's story into the era that followed the *Origin of Species*. Secondly, its focus, limited to a shorter and more recent period of time, shows us a not-distant mirror of our current disagreements (punctuated equilibria, neutralism, cladism, etc.). By implication it suggests that our question should not be "Which?" but "How much of each?" □

Barry Cox is Head of the Zoology Department at King's College, University of London.

Group intelligence

Richard Lynn

Handbook of Human Intelligence.

Edited by Robert J. Sternberg.

Cambridge University Press: 1983.

Pp. 1,025. Hbk £45, \$65; pbk £15, \$24.95.

CAMBRIDGE University Press present this book as the most comprehensive account ever published of human intelligence, one which will guide thinking about intelligence for the rest of the century. Such strident claims challenge the reader to disagree. However, there is no doubt that the editor, Dr Robert Sternberg of Yale, has succeeded in assembling a highly professional group of writers who have in general produced a thorough review volume in spite of some notable deficiencies.

The book is divided into five sections, the first of which deals with the nature of intelligence and its measurement. In the opening chapter Sternberg and William Salter tackle different conceptions and definitions of intelligence, concluding that the most common view has been of intelligence as an adaptive and goal-directed ability. There follows an account by John B. Carroll of the history and present state of the measurement of intelligence, with comments on the status of the general factor, the hereditary-environment issue and the politicization which has increasingly characterized the subject in recent decades.

Part Two of the book, entitled "Cognition, Personality and Intelligence", begins with a chapter by Lynn Cooper and Dennis Regan on attention, perception and intelligence. The authors conclude that there is little association between intelligence and low-level attentional and perceptual processes. This, however, has long been the established position and it is a pity that the chapter was apparently written before the authors had an opportunity to appraise the recent work of C. Brand calling this view into question. There follow contributions by William Estes on learning, memory and intelligence and by Sternberg on reasoning, problem solving and intelligence, this last chapter containing a good account of the information-processing approach to which Sternberg has made considerable contributions. This section of the book also contains chapters on personality and intelligence by Jonathan Baron, artificial intelligence and its relation to human intelligence by Natalie Dehn and Roger Schank, and mental retardation by Joseph Campione and colleagues.

Part Three, of three chapters, covers society, culture and intelligence. Richard Snow first deals with education and intelligence, then Edward Zigler and Victoria Seitz with social policy and intelligence — the authors are generally optimistic on the benefits of social intervention. The third chapter is on culture and intelligence, a col-

lective effort from the Laboratory of Comparative Human Cognition at the University of California. A common theme in this part of the book is the distinction between competence and performance and the social context within which intelligence operates.

Part Four is entitled "The Phylogeny and Ontogeny of Intelligence". In the first chapter Harry Jerison summarizes — but without significant addition — his brilliant reconstruction of the evolution of intelligence and the crucial role played by the demands of new niches in advancing intelligence. Next Sandra Scarr and Louise Carter-Saltzman review the heredity-environment issue, being moderately hereditarian on individual differences in intelligence but environmentalist on race differences. They summarize their recent and thought-provoking results on the race problem, for example the absence of any association among blacks between IQ and caucasian ancestry. The third chapter covers the development of intelligence in childhood, where Robert Siegler and Dean Richards provide a useful review of the approach of Piaget, information-processing theory and learning skills theory. The volume concludes with another chapter by Sternberg, this time with Janet Powell.

While this handbook provides thorough coverage of many salient areas of the subject, there are three weaknesses. First, there is no chapter on the social value of

ADVERTISEMENT

"OUTSTANDING"

THE ATOMIC COMPLEX

by BERTRAND GOLDSCHMIDT

An accurate, complete and fascinating worldwide political history and personal memoir of nuclear energy — from the development of the bomb in World War II to today's nuclear energy complex and proliferation problems.

"READS EASILY"

Goldschmidt, a leading French scientist turned international statesman, reviews half a century of political moves, countermoves, international intrigue and manipulation. *The Atomic Complex* is a pragmatic look at the nuclear world today, carefully examining the terror nuclear weapons represent, but at the same time stressing the benefits of nuclear energy.

Order this 500 page
best seller today.

\$31 Hardbound \$24 Softbound

**American Nuclear Society
555 North Kensington Avenue
La Grange Park, IL 60525 USA**

Circle No.04 on Reader Service Card.

intelligence. Francis Galton in the last century argued that high population intelligence was an important requirement for the achievement of a civilization with high economic, cultural and intellectual attainments. This theme was re-stated in the 1930s by Cyril Burt and E.L. Thorndike and more recently by R.B. Cattell and John Baker. It was belief in the social value of intelligence that gave rise to concern that dysgenic trends might be operating in advanced societies, with the consequent decline of national levels of intelligence. It might have been expected that a handbook aiming at comprehensive coverage would have provided some discussion of these important questions.

A second omission concerns the physiology of intelligence. The discovery during the last decade of the high correlation between IQ and the EEG evoked potential promises to be one of the significant developments in our understanding of the brain mechanisms underlying intelligence; it is a pity that there is no mention of this work.

Finally, and not trivially, the index is totally inadequate. An index normally runs to about two pages per one hundred pages of text. In this volume the index consists of barely four pages to service a thousand pages of text. The result is that there is no appearance in the index of many important topics discussed in the text, even including *g* (general intelligence), arguably the most important concept ever formulated in this field. It is a matter for astonishment that the publishers and editor — no doubt people of considerable ability — are apparently not aware that a handbook must have a good index if it is to serve its purpose. Such are the enigmas of intelligence. □

Richard Lynn is Professor of Psychology at the New University of Ulster, Coleraine.

Sea changes

A.J. Southward

Climate and Fisheries.

By D.H. Cushing.

Academic: 1983. Pp.372. £28, \$49.50.

HISTORY shows how the fortunes of towns and states rose and fell with the success or failure of the great northern fisheries, such as cod and herring. Many reasons have been advanced for the variations in abundance and occurrence of the shoals, but changes in weather and climate have long been the favourite, and fishery scientists have to spend much time assessing fluctuations in sea temperature, water currents and winds. Dr Cushing is particularly fitted to write on this subject; he is well known for his studies of the impact of recent climatic changes on marine life, and is also an expert on production processes in the sea, especially the links between algal

growth, herbivore grazing and young fish survival, one of the mechanisms controlling the abundance of the shoals.

The first four chapters introduce the subject: a brief summary of the major fisheries and their technologies; a survey of the way in which different fishes spawn, grow and migrate; a discussion of past changes; and a review of the meteorological background to climatic change. The author then moves deeper into his subject to describe the advance north of invertebrates and fish during the warm period between 1920 and 1960. This is followed with a review of all aspects of short-term fluctuation in fish stocks, which the author would evidently like to relate to year-to-year differences in phytoplankton production. There are other chapters dealing in detail with changes in the North Sea and the Baltic, perhaps the best studied seas in the world, and a description of the effects of very cold winters.

Dr Cushing devotes a whole chapter to the sorry tale of the Peruvian anchoveta fishery, which collapsed in 1972 after heavy fishing during a low point in a natural fluctuation. He notes how researches since then have increased our knowledge of hydrographic events in the Pacific to the point where it is now possible to predict intensification of the south-going warm current (the so-called El Niño) that drives the fish away and reduces recruitment. However he does not rub home sufficiently the message that continued overfishing contributed to the collapse, as happened with other fish that vanished during a period of natural decline, such as the California sardine.

Of course, the author's principal concern, as he stresses, is with natural changes. In the final chapters he discusses the mechanisms of natural change and the consequences from the point of view of fishery management. Many of the world's fisheries are now in a state of decline due to over-exploitation. It is vitally important that fishery managers should be able to distinguish between natural fluctuations and man-made changes, and assess the risks of allowing continued fishing on a smaller and fluctuating stock.

There are a number of irritating minor blemishes, including persistent misspellings, misunderstandings of other authors' work and garbled paragraphs, suggesting failure to consult primary sources and lack of critical reading of the manuscript. In addition, the author's terse style makes little concession to the non-specialist reader, and beginners must look elsewhere for enlightenment. These, however, are relatively minor drawbacks. The book provides a fact-filled summary that is essential for all marine scientists, and is a must for those concerned with fishery management, environmental assessment and political economics. □

A.J. Southward is at the Marine Biological Association's Laboratory, Plymouth.

Neuroscience meets immunology

Ron McKay

Neuroimmunology.

Edited by Jeremy Brockes.

Plenum: 1982. Pp.256. \$29.50, £20.65.

THERE have been two major technical developments in molecular biology in recent years — recombinant DNA and hybridoma technology. Both of these techniques are only just beginning to be applied to the outstanding problems in neurobiology, but already they promise to revolutionize our understanding in many areas of neuroscience.

In this book, Jeremy Brockes has gathered eight contributions which discuss the use of immunological techniques to study the nervous system. The primary reason for using such procedures is the exquisite specificity with which antibody can bind to antigen. However, until recently the complex array of antibodies produced in an immune response was the main difficulty in applying immunoglobulins (which bind specifically to antigens) to probe biological structures. The solution to the problem was the experiment of Kohler and Milstein, which showed that lymphocytes secreting different antibodies could be rescued from a rapid death in culture if they were fused with myeloma cells and could then be separated to provide access to the component antibodies. This experiment, which grew out of a marriage between immunology and somatic cell genetics, gave us fresh access to the large number of specific antibodies synthesized by the vertebrate immune system.

There are many applications of immunological technology to studies of the nervous system. In this book two chapters are devoted to the use of antibodies in identifying and characterizing the components of the neuromuscular junction. A third chapter discusses the application of monoclonal antibodies to map synaptic constituents of the vertebrate central nervous system. Three further contributions consider the use of monoclonal antibodies in defining the cellular heterogeneity among neurons and glia, these chapters being buttressed by an account of the strikingly successful use of specific immunoglobulins to identify, purify and functionally define subsets of cells in the immune system. Finally, the pathologic interaction of the immune system and the nervous system is considered. As indicated by the variety of the contributions in the book, the term neuroimmunology itself is vague in meaning; but common to all of the chapters is the sense of excitement over the early successes of this means of identifying specific molecules in a complex tissue which is resistant to standard biochemical techniques.

The neuromuscular junction (NMJ) is a

favoured system for studying synaptic interaction because the two cell types involved are clearly defined. The most striking results of the immunological approach in this area are twofold: first, the identification of previously unrecognized molecular components of the neuromuscular junction and, second, the analysis of the structure and function of the acetylcholine receptor (for example, some monoclonal antibodies have been reported which block the binding of alpha-toxins to the receptor and others which block sodium efflux). Our new ability to identify new molecules with monoclonal antibodies and then study in detail the contribution of different sites to the function of the molecule is clearly illustrated in the book.

Rather than studying the simple cellular system of the NMJ, several groups are now using hybridoma technology to dissect the cellular complexity of the central nervous system. Antibodies are available which define the major cell types of the nervous system, neurons and glia. It is becoming increasingly clear from immunological studies that the nervous systems of both

vertebrates and invertebrates are composed of many antigenically distinct classes of cells. While there have been several different lines of evidence which suggested the nervous system was complex at the cellular level, hybridoma technology offers us a general method for establishing the extent of this cellular diversity. Of course these investigators are not only interested in the number of cell types, but also in how they interact, where they come from in development and, ultimately, which molecules are responsible for organizing important features of the nervous system.

Neuroimmunology provides a lucid and scholarly introduction to this area of research. The limitations as well as the advantages of the techniques are clearly expressed. While the term neuroimmunology will very probably fade in time, the techniques described here and the insights they give us will become a standard part of our understanding of the molecular basis of neuronal organization. □

Ron McKay is at Cold Spring Harbor Laboratory, Long Island, New York.

The Earth's fluid envelope

James Lighthill

Atmosphere–Ocean Dynamics.

By Adrian E. Gill.

Academic: 1982. Pp. 660.

Hbk \$60, £39.60; pbk \$30, £19.80.

FOR the earth sciences as a whole, the appearance of Adrian Gill's book is an event of very special importance. It is the first work to give comprehensive and mutually integrated descriptions of the atmosphere and ocean, those two complex and intimately interacting dynamical systems that modify so greatly our planet's response to the distribution of incoming solar energy.

Those dynamical systems are both fluid media; both are strongly influenced by the Earth's gravity, its rotation and its solid-surface topography, as well as by density stratification and by mutual interaction. Many authors have drawn attention to the wide range of closely analogous dynamical properties exhibited by the two systems, while others have chosen to emphasize the differences between them associated with the contrasting physical properties of salt water and moist air. Here is a book which combines those two approaches. Equally, Dr Gill is the first author to treat the interaction between atmosphere and ocean in a degree of detail responsive to the complexity of both systems. He also makes good use of the results of numerical-modelling studies. At the same time, the book shows how a full knowledge of those various fundamental features of a fluid's

motion, exhibited on various scales (dependent on the physical properties of the medium) by media subject to gravity, stratification, rotation and different kinds of external forcing, is essential — whether for effective utilization of the results of numerical modelling, or for estimating the modelling grid-scales required for particular purposes.

The book begins with indications of what the fluid dynamics will have to explain. First, a thorough account of radiation balance in the atmosphere/ocean/land system exhibits the major excess at low latitudes (compensated by a deficit at high latitudes) of solar-energy absorption over outgoing radiation. Next, the vital need to see atmosphere and ocean as a single system is emphasized by another data summary (Figure 1.8) indicating how poleward energy flow is shared about equally between the two. The force balance at the atmosphere's base is then shown to demand an atmospheric eddy transport of angular momentum up its gradient in mid-latitudes. In addition, the complex dependence of drag coefficients at the air-sea interface on wave height and hence on wind intensity is discussed, followed by a preliminary view of wind-stress and thermohaline forcing as the two main determinants (in that order) of ocean circulation. Finally, the hydrological cycle is depicted in detail, with full accounts of all thermodynamic and hydrostatic-stability analysis relevant to the Earth's fluid envelope from the ocean deeps to the

stratosphere and mesosphere.

In the subsequent, dynamical chapters, Dr Gill demonstrates the extraordinary degree of commonality of fundamental dynamical features of both systems (the same feature being exhibited in different ranges of length-scale in each). He uses this commonality to achieve a most satisfyingly unified presentation of those essential fluid-dynamical features that can be used so effectively to interpret data when combined with the results of numerical modelling.

Preliminary material on features not significantly affected by Earth's rotation covers the special properties of highly dispersive gravity waves, such as surface and internal waves, and weakly dispersive waves, such as long waves. This material includes extensive application to seiches, tides and undercurrents in narrow channels, and to waves in the ocean thermocline; to that large component of wind drag over land provided by the wave-drag associated with hills' lee-wave patterns; and to the atmospheric propagation of waves from big explosions.

Intermediate material on those larger-scale phenomena that are strongly influenced by Earth's rotation begins by demonstrating the crucial importance of the scale provided by the Rossby radius of deformation. Phenomena on greater length-scales are dominated by geostrophic adjustment (with its implications on how the component of fluid motion along isotherms increases with height, and its fruitful interpretations in terms of potential vorticity). More accurately, they can be described on a quasi-geostrophic approximation in which the geostrophic motion along isobars is accompanied by another motion perpendicular to the isallobars (the divergence of the isallobaric component is an important determinant of vertical motions). At the same time, phenomena depending on the effects of Earth's rotation on internal waves and on long waves are described in detail, together with their important limiting forms near the ends of the relevant frequency ranges. The book includes a particularly impressive account of six types of internal-wave generation by fluid flow over topography, each type being found in an appropriate range of length-scales. Wave "ray tracing", as well as wave absorption (especially in critical layers) and wave dissipation, are also comprehensively treated.

Another important part of the book deals with forcing of the atmosphere and ocean by three factors — tide-raising effects; momentum, heat and water exchange at their interface; and side-boundary effects. Ekman transport in both atmosphere and ocean (the response to their mutual stress interaction) is admirably discussed, along with the "pumping" that results when that transport has non-zero divergence. An excellent treatment of coastally trapped long waves in the ocean (including Kelvin waves and shelf

● Next week's issue of *Nature* includes the Spring Books Supplement. Some 30 books are reviewed, including Derek Freeman's *Margaret Mead and Samoa*, Roger Sperry's *Science and Moral Priority* and Hideki Yukawa's *Tabibito*.

waves) is given alongside a good general account of tides and storm surges. Coastal upwelling, too, is lucidly analysed, and this part of the book ends with a list of 16 penetrating comments on features that contribute to the puzzling nature of Eastern boundary currents.

Two chapters which follow indicate the special features of the atmosphere-ocean system in the tropics and in mid-latitudes. The ocean's equatorially trapped waves are analysed, along with equatorial undercurrents. The important vertically propagating waves in the equatorial atmosphere are related — like all the dynamical features covered in the book — to clear observational data establishing their significance. The special characteristics of the SW Monsoon winds of the Indian Ocean and the variable ocean currents which they generate are excellently treated. The very different characteristics of mid-latitude motions are seen as in part due to the enormous frequency separation between internal waves and planetary waves. Consequences for the time-scale of an ocean's response to changing wind stress are analysed and related to the characteristics of Western boundary currents such as the Gulf Stream and the Kuroshio. This is another section where the quasi-geostrophic approximation is found to be most fruitful. In the atmosphere, the Eliassen-Palm Flux of energy, in the direction of a certain wave group velocity, is found to be a most powerful mechanism for feeding the jet stream.

The last chapter outlines the crucial large-scale instabilities that are needed in the atmosphere to provide the angular momentum flux which the general circulation demands. The author's description of the crucial barotropic and baroclinic instabilities is particularly well done, and leads to the outline of the life-cycle of a baroclinic disturbance such as a mid-latitude cyclone or (say) one of the ocean eddies that peel off from the Gulf Stream. Frontogenesis is described both in the atmosphere and in the ocean. Finally, the Eliassen-Palm Flux is seen as generating simultaneously the required transport of energy and of angular momentum consistent with the latitudinal distribution of surface winds.

In all, Dr Gill has demonstrated the great value of an exposition on this subject by an author equally proficient in both meteorology and physical oceanography. His book will be indispensable to those who would like to learn about both the extremely complicated, observed data on the atmosphere and ocean, and the ways in which modern fluid dynamics can be used to illuminate them. Those ways are most impressive even though they are still incomplete. As Dr Gill observes in his final sentence, "nature is complex and there is much to be learned!" □

Sir James Lighthill is Provost of University College London.

Are sons worth more than daughters?

Kenneth B. Armitage

Red Deer: Behaviour and Ecology of Two Sexes.

By T.H. Clutton-Brock, F.E. Guinness and S.D. Albon.

Edinburgh University Press/University of Chicago Press: 1983. Pp.400. Hbk £20; pbk £9.75, \$12.95.

A RECENT article in *Science* 83 asked why men are bigger than women. The answers provided were speculative and characteristic of much contemporary writing in sociobiology and evolutionary biology in which every difference, every behaviour, has an *a posteriori* adaptive explanation. Not all the pitfalls of speculative interpretation are avoided in this book, but overall it represents a significant advance in the emerging sub-discipline of behavioural ecology.

The authors sought to "examine the evolutionary causes and ecological consequence of sex difference" in the red deer, especially that of size and reproduction. Each individual attempts to maximize its reproductive success. The behaviour, ecology and consequences associated with this strategy not only reflect differences among individuals, but also differences between sexes.

In the studies described in the book, the variance in lifetime reproductive success was greater among males than females. Hinds invested more in their sons before weaning; sons are born heavier, require a longer gestation period and suckle more frequently than do daughters. As evolutionary theory predicts that, overall, investment will be equally divided between sons and daughters, it was expected that slightly fewer of the more "expensive" sons would be produced — there should be a slight bias towards females in the sex ratio. As it turned out, the bias was towards males. To reconcile observations and theory, the authors suggest that investment was equalized by a greater post-weaning investment in daughters. Daughters stay with their mothers after weaning and use the same feeding area, apparently at a cost to the mother.

Looking at investment in sons or daughters in this way does not, however, take into account the longer term pay-offs. Increased direct fitness requires grand offspring and the fitness pay-off for one successful son — who can fertilize many females — may equal that of several daughters. Before we can formulate appropriate models of sex ratios we need to know the probabilities of producing reproductively successful sons and daughters. In general, however, the book lacks a thorough treatment of kinship and of its relationship to reproductive success, to the movement of hinds between harems and to

population growth. Kinship needs to be understood before coming to the conclusion that scramble competition for resources occurs among females. Female groups are matrilineal; if resources are too widely distributed to be easily defended, fitness could be enhanced by sharing resources with kin. Is it the sharing of resources that leads to the formation of groups rather than danger from predators? If that is the case, eventually population increase could lead to competition among kin. How will that competition be expressed?

I especially enjoyed the chain of relationships demonstrated repeatedly throughout the book. For example, although fighting among stags is costly, fighting ability determines harem size and reproductive success; body size, a consequence of early growth, in turn determines fighting ability. In cold winters, however, large size may be detrimental, male mortality exceeding that of females. Does winter stress act as a counter-

IMAGE
UNAVAILABLE FOR
COPYRIGHT
REASONS

selection against increasing male size? Again, variation in the lifetime reproductive success of hinds was more a consequence of calf survival than of fecundity. Hinds did not select particular stags as mates other than avoiding immature stags. Birth weight of calves affected summer survival, but birth date was a better predictor of winter survival.

Although it was pointed out a decade ago that male and female strategies may differ and even conflict, this book makes a major contribution in thoroughly documenting the importance of treating each sex separately when analysing the evolutionary ecology of a species. Further, the authors have clearly demonstrated that a long-term research effort will be needed in order to replace adaptationist stories with rigorous statistical analysis of competing hypotheses. □

Kenneth B. Armitage is Professor of Systematics and Ecology at the University of Kansas.